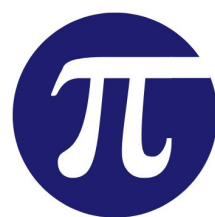




PROGRAD
Pró-Reitoria de
Graduação



PETIMAT
Programa de Educação Tutorial
Institucional de Matemática

Notas de estudo

Um estudo numérico do método dos gradientes conjugados projetados

Wéllington Moutinho Dias

Conteúdo

1	Revisão de Álgebra Linear	2
2	Introdução	5
3	Algoritmo	5
4	Análise de Convergência	14
5	Taxa de Convergência	19
6	Experimentos Numéricos	20

1 Revisão de Álgebra Linear

Definição 1.1. Dados, $m, n \in \mathbb{N}$. Uma $m \times n$ **matriz** sobre \mathbb{R} é uma função $A : I_m \times I_n \rightarrow \mathbb{R}$ tal que $A(i, j) = a_{ij}$ para todo $(i, j) \in I_m \times I_n$. Onde $I_n = \{p \in \mathbb{N} : p \leq n\}$.

Quando $m = n$, dizemos que uma $m \times n$ matriz sobre \mathbb{R} é quadrada. Usualmente, os valores que uma $m \times n$ **matriz** sobre \mathbb{R} , digamos A , assume são representados por uma tabela retangular e denotados da seguinte maneira

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

a definição a seguir refere-se a tal representação.

Definição 1.2. Uma **operação elementar sobre as linhas** de uma $m \times n$ A sobre \mathbb{R} refere-se a

- (a) multiplicação de uma linha de A por um escalar c não-nulo;
- (b) substituição da r -ésima linha de A pela linha r mais c vezes a linha s , sendo c um escalar arbitrário e $r \neq s$;
- (c) transposição de duas linhas de A .

Definição 1.3. Uma $m \times n$ matriz R é dita **escalonada reduzida** se

- (a) o primeiro elemento não-nulo em cada linha não-nula de R é igual a 1;
- (b) cada coluna de R que contém o primeiro elemento não-nulo de alguma linha tem todos os seus outros elementos nulos;
- (c) toda linha de R cujos elementos são todos nulos ocorre abaixo de todas as linhas que possuem um elemento não-nulo;
- (d) se as linhas $1, \dots, r$ são as linhas não-nulas de R e se o primeiro elemento não-nulo da linha i ocorre na coluna k_i , $i = 1, \dots, r$, então $k_1 < k_2 < \dots < k_r$.

Definição 1.4. O **posto** de uma matriz A é o número de linhas não nulas da sua forma escalonada reduzida.

Como não trabalharemos com matrizes sobre outros corpos e também não teremos o perigo de confundir seu “tamanho” diremos apenas o termo “matriz” ao invés de $m \times n$ matriz sobre \mathbb{R} .

Definição 1.5. Dada uma matriz A de tamanho $m \times n$, a $n \times m$ matriz **transposta** denotada por A^\top é uma função $A^\top : I_n \times I_m \rightarrow \mathbb{R}$ tal que $A^\top(i, j) = a_{ji}$ para todo $(i, j) \in I_n \times I_m$.

A transposição de matrizes possuem algumas propriedades algébricas, enunciaremos as que serão utilizadas em nosso estudo e omitiremos suas demonstrações.

Proposição 1.1. Dada uma matriz A temos que $(A^\top)^\top = A$.

Definição 1.6. Dadas uma matriz $m \times n$ A e uma matriz $n \times p$ B . Definiremos a $m \times p$ matriz produto $C = A \cdot B$ de modo que $C(i, j) = \sum_{k=1}^n a_{ik} \cdot b_{kj}$.

Há uma propriedade algébrica que relaciona o produto de matrizes com a transposição de matrizes.

Proposição 1.2. Dadas uma matriz $m \times n$ A e uma matriz $n \times p$ B . Temos que

$$(AB)^\top = B^\top \cdot A^\top.$$

Definição 1.7. Dadas as matrizes $m \times n$ A e B . Definiremos a $m \times n$ matriz soma $C = A + B$ de modo que $C(i, j) = A(i, j) + B(i, j)$.

Proposição 1.3. Dadas as matrizes $m \times n$ A e B temos que $(A + B)^\top = A^\top + B^\top$.

Definição 1.8. Uma matriz A é dita **simétrica** se $A^\top = A$.

Proposição 1.4. A soma de matrizes simétricas é simétrica.

Demonstração. Sejam A e B matrizes simétricas. Vamos mostrar que $A + B$ também é simétrica. Para isso, observe que

$$(A + B)^\top = A^\top + B^\top = A + B.$$

□

Os espaços vetoriais \mathbb{R}^n e $\mathbb{R}^{n \times 1}$ (espaço das matrizes $n \times 1$ sobre \mathbb{R}) são isomorfos logo, o produto de matrizes definido acima também é válido para os elementos do \mathbb{R}^n . Precisaremos dessa informação para resolver a proposição a seguir.

Proposição 1.5. Dados $u, v \in \mathbb{R}^n \setminus \{0\}$, a matriz gerada por uv^\top tem posto 1.

Demonstração. Sejam $u = (u_1, u_2, \dots, u_n)$ e $v = (v_1, v_2, \dots, v_n)$, assim

$$uv^\top = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} [v_1 \ v_2 \ \dots \ v_n] = \begin{bmatrix} u_1v_1 & u_1v_2 & \dots & u_1v_n \\ u_2v_1 & u_2v_2 & \dots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nv_1 & u_nv_2 & \dots & u_nv_n \end{bmatrix} \stackrel{(*)}{\sim} \begin{bmatrix} u_1v_1 & u_1v_2 & \dots & u_1v_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Como $u \neq 0$, então existe um $i \in \{1, 2, \dots, n\}$ tal que $u_i \neq 0$, assim a matriz obtida por $(*)$ é resultante das seguintes operações em linhas,

$l_j = -\frac{u_i}{u_j}l_j + l_i, i \neq j, u_j \neq 0$ (substituição da j -ésima linha por $-\frac{u_i}{u_j}$ vezes a j -ésima linha mais a i -ésima linha com $i \neq j$ e $u_j \neq 0$);

$l_i \Leftrightarrow l_1$ (seguida da transposição da i -ésima linha à primeira linha).

Observe que se $u_j = 0$ então a j -ésima linha da matriz é nula. Portanto, uv^\top tem posto 1. □

Em seguida enunciaremos alguns resultados cuja demonstração podem ser encontradas em vários livros de álgebra linear.

Proposição 1.6. Para qualquer matriz A , $\text{posto}(A^\top) = \text{posto}(A)$.

Proposição 1.7. Para quaisquer $A, B \in \mathbb{R}^{n \times m}$, $\text{posto}(A + B) \leq \text{posto}(A) + \text{posto}(B)$.

Corolário 1.1. Dados $u, v \in \mathbb{R}^n \setminus \{0\}$, $\text{posto}(uv^\top + vu^\top) \leq 2$.

Demonstração. Com efeito,

$$\begin{aligned} \text{posto}(uv^\top + vu^\top) &\leq \text{posto}(uv^\top) + \text{posto}(vu^\top) \text{ (pela Proposição 1.7)} \\ &= \text{posto}(uv^\top) + \text{posto}((vu^\top)^\top) \text{ (pela Proposição 1.6)} \\ &= \text{posto}(uv^\top) + \text{posto}((u^\top)^\top \cdot v^\top) \text{ (pela Proposição 1.2)} \\ &= \text{posto}(uv^\top) + \text{posto}(uv^\top) \text{ (pela Proposição 1.1)} \\ &= 2 \cdot \text{posto}(uv^\top) \\ &= 2 \text{ (pela Proposição 1.5)}. \end{aligned}$$

□

Vamos relembrar sobre uma matriz especial. Denominamos por **matriz identidade** e denotamos por I_n a matriz quadrada $n \times n$ tal que

$$I_n(i, j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases},$$

ou seja, sua representação tabular é dada por

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

O conjunto das matrizes quadradas $n \times n$ sobre \mathbb{R} denotado por $\mathbb{R}^{n \times n}$ munido do produto definido anteriormente possui como elemento a matriz identidade definida acima. No entanto, nem todas as matrizes $A \in \mathbb{R}^{n \times n}$ possuem um $B \in \mathbb{R}^{n \times n}$ de modo que $A \cdot B = B \cdot A = I_n$, isto é, que possuem inverso multiplicativo. Apesar disso, as matrizes que possuem recebem um nome especial.

Definição 1.9. Uma matriz $A \in \mathbb{R}^{n \times n}$ é **invertível** se existir um $B \in \mathbb{R}^{n \times n}$ tal que $A \cdot B = B \cdot A = I_n$.

É fácil mostrar a unicidade da matriz B na definição dada acima. Por isso, a matriz B é denotada por A^{-1} e é chamada de **matriz inversa de A**.

Definição 1.10. Sejam A e B matrizes quadradas. Dizemos que A é **semelhante** a B se existir uma matriz invertível P tal que $P^{-1}AP = B$. Se A é semelhante a B , escreveremos $A \sim B$.

Definição 1.11. Uma matriz quadrada A é dita **diagonal** se $A(i, j) = 0$ caso $i \neq j$.

Observe que a matriz identidade é um exemplo de uma matriz diagonal.

Definição 1.12. Uma matriz A é **diagonalizável** se $A \sim D$, com D sendo uma matriz diagonal.

Ou seja, pelas definições dadas acima, temos que A é diagonalizável é equivalente a afirmar que existe uma matriz invertível P tal que $P^{-1}AP = D$ com D sendo uma matriz diagonal. A proposição a seguir, cuja demonstração será omitida, descreve os elementos de cada coluna da matriz P e os elementos não nulos da matriz diagonal D . Para isso, precisamos recordar os seguintes conceitos.

Definição 1.13. Sejam $A \in \mathbb{R}^{n \times m}$ e $v \in \mathbb{R}^m \setminus \{0\}$. Dizemos que v é um **autovetor associado ao autovalor** $\lambda \in \mathbb{C}$ se $A \cdot v = \lambda \cdot v$.

Proposição 1.8. Seja A uma matriz $n \times n$. Existem uma matriz invertível P e uma matriz diagonal D de maneira que $P^{-1}AP = D$ se e somente se as colunas de P forem n autovetores de A , linearmente independentes, e os elementos da diagonal principal de D forem os autovalores correspondentes aos autovetores em P , colocados na mesma ordem.

Definição 1.14. Um produto interno em \mathbb{R}^n é uma função $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ que satisfaz as seguintes propriedades para quaisquer $u, v, w \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$:

positividade $\langle u, u \rangle \geq 0$;

produto interno nulo $\langle u, u \rangle = 0$ se, e somente se, $u = 0$;

linearidade na primeira entrada $\langle \alpha u + v, w \rangle = \alpha \langle u, w \rangle + \langle v, w \rangle$;

simetria $\langle u, v \rangle = \langle v, u \rangle$.

Dados $u = (u_1, u_2, \dots, u_n)$ e $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ a função $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $\langle u, v \rangle = u_1v_1 + u_2v_2 + \cdots + u_nv_n$ é um exemplo de produto interno. Mas note que $u^\top v = \langle u, v \rangle$ e pela simetria temos que $u^\top v = v^\top u$. Por fim, uma identidade recorrente é $u^\top u = \|u\|^2$.

Definição 1.15. Dizemos que os vetores $u, v \in \mathbb{R}^n$ são **ortogonais** se $\langle u, v \rangle = 0$.

Definição 1.16. Um conjunto de vetores de \mathbb{R}^n é um **conjunto ortogonal** se todos os pares de vetores distintos no conjunto são ortogonais.

Definição 1.17. Uma norma em \mathbb{R}^n é uma função $\|\cdot\| : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $\|x\| = \sqrt{\langle x, x \rangle}$ para todo $x \in \mathbb{R}^n$.

Definição 1.18. Dizemos que $x \in \mathbb{R}^n$ é **unitário** se $\|x\| = 1$.

Definição 1.19. Um conjunto de vetores de \mathbb{R}^n é um **conjunto ortonormal** se todos os pares de vetores distintos no conjunto são ortogonais e unitários.

Definição 1.20. Uma matriz quadrada Q cujas colunas formam um conjunto ortonormal é chamada **matriz ortogonal**.

Definição 1.21. Uma matriz quadrada A é **ortogonalmente diagonalizável** se existirem uma matriz ortogonal Q e uma matriz diagonal D tais que $Q^T A Q = D$.

Uma consequência direta da definição acima é que toda matriz ortogonalmente diagonalizável é também diagonalizável.

Teorema 1.1 (espectral). *Seja A uma matriz quadrada. Então, A é simétrica se e somente se A é ortogonalmente diagonalizável.*

Proposição 1.9. *Sejam A e B matrizes quadradas com $A \sim B$. Então A e B tem o mesmo posto.*

Proposição 1.10. *Sejam A uma matriz quadrada e v um autovetor. Se λ é um autovalor associado a v , então $\beta + \gamma\alpha$ é um autovalor associado a $\beta \cdot I_n + \gamma \cdot A$, isto é*

$$(\beta \cdot I_n + \gamma \cdot A)v = (\beta + \gamma\alpha)v.$$

Proposição 1.11. *Para quaisquer vetores $u, v \in \mathbb{R}^n$, temos que, $(uv^T)u = u(v^T u)$.*

Demonstração. De fato,

$$\begin{aligned} (uv^T)u &= \begin{pmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} & \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \end{pmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \\ &= \begin{bmatrix} u_1 v_1 & u_1 v_2 & \dots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \dots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n v_1 & u_n v_2 & \dots & u_n v_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \\ &= \begin{bmatrix} u_1^2 v_1 + u_1 u_2 v_2 + \dots + u_1 u_n v_n \\ u_1 u_2 v_1 + u_2^2 v_2 + \dots + u_2 u_n v_n \\ \vdots \\ u_1 u_n v_1 + u_2 u_n v_2 + \dots + u_n^2 v_n \end{bmatrix} \\ &= \begin{bmatrix} u_1(u_1 v_1 + u_2 v_2 + \dots + u_n v_n) \\ u_2(u_1 v_1 + u_2 v_2 + \dots + u_n v_n) \\ \vdots \\ u_n(u_1 v_1 + u_2 v_2 + \dots + u_n v_n) \end{bmatrix} \\ &= \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} (u_1 v_1 + u_2 v_2 + \dots + u_n v_n) \\ &= \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} (v_1 u_1 + v_2 u_2 + \dots + v_n u_n) \\ &= u(v^T u). \end{aligned}$$

□

Proposição 1.12. *Para quaisquer vetores $u \in \mathbb{R}^n$ e $v \in \mathbb{R}^n \setminus \{0\}$, temos que,*

$$\left(uv^T + vu^T \right) \left(u \pm \frac{\|u\|}{\|v\|} v \right) = \left(u^T v \pm \|u\| \|v\| \right) \left(u \pm \frac{\|u\|}{\|v\|} v \right).$$

Demonstração. De fato,

$$\begin{aligned} \left(uv^T + vu^T \right) \left(u + \frac{\|u\|}{\|v\|} v \right) &= (uv^T)u + (vu^T)u + \frac{\|u\|}{\|v\|} (uv^T)v + \frac{\|u\|}{\|v\|} (vu^T)v \\ &= u(v^T u) + v(u^T u) + \frac{\|u\|}{\|v\|} u(v^T v) + \frac{\|u\|}{\|v\|} v(u^T v) \\ &= u(u^T v) + \frac{\|u\|}{\|v\|} v(u^T v) + \|u\|^2 v + \frac{\|u\|}{\|v\|} \|v\|^2 u \\ &= \left(u + \frac{\|u\|}{\|v\|} v \right) (u^T v) + \|u\| \|v\| \left(\frac{\|u\|}{\|v\|} v \right) + \|u\| \|v\| u \\ &= \left(u + \frac{\|u\|}{\|v\|} v \right) (u^T v) + \left(u + \frac{\|u\|}{\|v\|} v \right) \|u\| \|v\| \\ &= (u^T v + \|u\| \|v\|) \left(u + \frac{\|u\|}{\|v\|} v \right). \end{aligned}$$

Além disso,

$$\begin{aligned}
\left(uv^\top + vu^\top\right)\left(u - \frac{\|u\|}{\|v\|}v\right) &= (uv^\top)u + (vu^\top)u - \frac{\|u\|}{\|v\|}(uv^\top)v - \frac{\|u\|}{\|v\|}(vu^\top)v \\
&= u(v^\top u) + v(u^\top u) - \frac{\|u\|}{\|v\|}u(v^\top v) - \frac{\|u\|}{\|v\|}v(u^\top v) \\
&= u(u^\top v) - \frac{\|u\|}{\|v\|}v(u^\top v) + \|u\|^2v - \frac{\|u\|}{\|v\|}\|v\|^2u \\
&= \left(u - \frac{\|u\|}{\|v\|}v\right)(u^\top v) + \|u\|\|v\|\left(\frac{\|u\|}{\|v\|}v\right) - \|u\|\|v\|u \\
&= \left(u - \frac{\|u\|}{\|v\|}v\right)(u^\top v) + \left(u - \frac{\|u\|}{\|v\|}v\right)(-\|u\|\|v\|) \\
&= (u^\top v - \|u\|\|v\|)\left(u - \frac{\|u\|}{\|v\|}v\right).
\end{aligned}$$

Portanto,

$$\left(uv^\top + vu^\top\right)\left(u \pm \frac{\|u\|}{\|v\|}v\right) = \left(u^\top v \pm \|u\|\|v\|\right)\left(u \pm \frac{\|u\|}{\|v\|}v\right).$$

□

Logo, se uma matriz quadrada pode ser decomposta na forma acima, então já sabemos pelo menos dois de seus autovalores.

Teorema 1.2. *Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica com λ_1 e λ_n , sendo o menor e o maior autovalor, respectivamente, então*

$$\lambda_1\|x\|^2 \leq x^\top Ax \leq \lambda_n\|x\|^2,$$

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

Proposição 1.13. *Dados $u, v, x \in \mathbb{R}^n$ temos que $x^\top uv^\top x = x^\top vu^\top x$.*

Demonstração. Repare que

$$\begin{aligned}
x^\top uv^\top x &= (x^\top u)(v^\top x) \text{ (associatividade em } \mathbb{R}\text{)} \\
&= (v^\top x)(x^\top u) \text{ (comutatividade em } \mathbb{R}\text{)} \\
&= (x^\top v)(u^\top x) \text{ (simetria do produto interno em } \mathbb{R}^n\text{)} \\
&= x^\top vu^\top x.
\end{aligned}$$

□

2 Introdução

Nesse estudo, os autores consideraram resolver a seguinte equação

$$F(x) = 0, \quad x \in \Omega, \quad (1)$$

onde $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é de classe C^1 e satisfaz a desigualdade abaixo para quaisquer $x, y \in \mathbb{R}^n$

$$(F(x) - F(y))^\top(x - y) \geq 0, \quad (2)$$

a propriedade acima será referida como a monotonicidade do campo vetorial F . O conjunto viável Ω é sempre não vazio, fechado e convexo.

Para resolver (1), os autores desenvolveram um método do gradiente conjugado projetado. Para isso, adaptaram o método do gradiente conjugado proposto por Dai & Kou que visa a minimização irrestrita de funções escalares convexas.

3 Algoritmo

Apresentaremos a seguir a adaptação proposta. O método do gradiente conjugado proposto no artigo estudado gera uma sequência $\{x_k\}$ da forma

$$x_{k+1} = x_k + \alpha_k d_k, \quad (3)$$

onde $\alpha_k > 0$ é o parâmetro responsável pelo comprimento de passo e d_k é uma sequência gerada pela fórmula de recorrência

$$d_k = \begin{cases} -F_0, & \text{se } k = 0, \\ -F_k + \beta_k^+(\tau_{k-1})d_{k-1}, & \text{se } k \geq 1, \end{cases} \quad (4)$$

onde $F_k = F(x_k)$ e

$$\beta_k^+(\tau_{k-1}) = \max \left\{ \beta_k(\tau_{k-1}), \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2} \right\}, \quad \eta \in [0, 1), \quad (5)$$

onde

$$\beta_k(\tau_{k-1}) = \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}}, \quad (6)$$

em (6) τ_{k-1} é definido como

$$\tau_{k-1}^A = \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \quad \text{e} \quad \tau_{k-1}^B = \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2}, \quad (7)$$

ou suas variantes com $\theta \in [0, 1]$,

$$\tau_{k-1} = \theta \tau_{k-1}^A + (1 - \theta) \tau_{k-1}^B. \quad (8)$$

Além disso, y_{k-1} presente em (6) é definido por $y_{k-1} = \gamma_{k-1} + \lambda_{k-1} t_{k-1} \|F_{k-1}\| d_{k-1}$ onde $\gamma_{k-1} = F_k - F_{k-1}$, λ_{k-1} é dado por

$$\lambda_{k-1} = 1 + \|F_{k-1}\|^{-1} \max \left\{ 0, -\frac{t_{k-1} \gamma_{k-1}^\top d_{k-1}}{\|t_{k-1} d_{k-1}\|^2} \right\}, \quad (9)$$

ademais, $t_{k-1} > 0$ e por fim, s_{k-1} em (6) é definido como $s_{k-1} = t_{k-1} d_{k-1}$.

Definiu-se ainda o operador projeção $P_\Omega : \mathbb{R}^n \rightarrow \Omega$ tal que

$$P_\Omega(x) = \arg \min \{ \|y - x\| : y \in \Omega \}, \quad \forall x \in \mathbb{R}^n.$$

Sabemos que $P_\Omega(x)$ é a projeção de x sobre o conjunto Ω . Além disso, se $x \in \Omega$, $P_\Omega(x) = x$. Uma propriedade mencionável de $P_\Omega[\cdot]$ é a sua não expansividade, isto é,

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

Baseado nestas informações, o seguinte algoritmo foi apresentado.

Algoritmo 1: Método do Gradiente Conjugado Projetado Ding, Xiao & Li

Passo 0 Escolha um chute inicial arbitrário $x_0 \in \Omega$ e as constantes $\xi \in (0, 1)$, $\sigma \in (0, 1)$, $\rho \in (0, 1)$, $\eta \in [0, 1]$, $\theta \in [0, 1]$, $\epsilon > 0$. Defina $k := 0$.

Passo 1 Pare se $\|F(x_k)\| < \epsilon$. Caso contrário, determine d_k por (4).

Passo 2 Encontre $t_k = \max\{\xi \rho^i : i = 0, 1, 2, \dots\}$, que satisfaça

$$-F(x_k + t_k d_k)^\top d_k \geq \sigma t_k \|d_k\|^2. \quad (10)$$

Defina $z_k := x_k + t_k d_k$.

Passo 3 Se $\|F(z_k)\| < \epsilon$, pare. Caso contrário, calcule

$$x_{k+1} = P_\Omega(x_k - \alpha_k F(z_k)),$$

onde $\alpha_k = \frac{F(z_k)^\top (x_k - z_k)}{\|F(z_k)\|^2}$.

Passo 4 Faça $k := k + 1$. Retorne ao **Passo 1**.

Agora, será mostrado uma propriedade da sequência definida em (4).

Lema 3.1. Seja $\{d_k\}$ a sequência gerada por (4) e τ_k definido por (8). Se $d_{k-1}^\top y_{k-1} \neq 0$, existe uma constante $c \in (0, 1)$, tal que

$$F_k^\top d_k \leq -c \|\mathcal{F}_k\|^2. \quad (11)$$

Demonstração. Consideraremos dois casos. No caso de $\beta(\tau_{k-1}) \geq \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2}$, então $\beta^+(\tau_{k-1}) = \beta(\tau_{k-1})$. Primeiramente, provaremos que

$$F_k^\top d_k \leq -\min \left\{ r_{k-1}, \frac{3}{4} \right\} \|F_k\|^2,$$

onde $r_{k-1} = \tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}}$.

Por conveniência, definimos

$$\begin{aligned} p_{k-1} &= \frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2}, \\ c_{k-1} &= \tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2}, \\ b_{k-1} &= y_{k-1} - c_{k-1} s_{k-1}, \\ H_k^P &= I_n - \frac{d_{k-1} b_{k-1}^\top}{d_{k-1}^\top y_{k-1}}. \end{aligned}$$

Como $s_k = t_k d_k$, afirmamos que

$$d_k = -H_k^P F_k.$$

Pois,

$$-H_k^P F_k = -\left(I_n - \frac{d_{k-1} b_{k-1}^\top}{d_{k-1}^\top y_{k-1}} \right) F_k = -F_k + \frac{d_{k-1} b_{k-1}^\top}{d_{k-1}^\top y_{k-1}} F_k.$$

Além disso,

$$\begin{aligned} \frac{d_{k-1} b_{k-1}^\top}{d_{k-1}^\top y_{k-1}} &= \frac{d_{k-1} (y_{k-1} - c_{k-1} s_{k-1})^\top}{d_{k-1}^\top y_{k-1}} \\ &= \frac{d_{k-1} (y_{k-1}^\top - c_{k-1} s_{k-1}^\top)}{d_{k-1}^\top y_{k-1}} \\ &= \frac{d_{k-1} y_{k-1}^\top}{d_{k-1}^\top y_{k-1}} - \frac{c_{k-1} d_{k-1} s_{k-1}^\top}{d_{k-1}^\top y_{k-1}} \\ &= \frac{d_{k-1} y_{k-1}^\top}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{d_{k-1} s_{k-1}^\top}{d_{k-1}^\top y_{k-1}} \\ &= d_{k-1} \left(\frac{y_{k-1}^\top}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{s_{k-1}^\top}{d_{k-1}^\top y_{k-1}} \right). \end{aligned}$$

Consequentemente,

$$\begin{aligned}
-H_k^P F_k &= -F_k + \frac{d_{k-1} b_{k-1}^\top}{d_{k-1}^\top y_{k-1}} F_k \\
&= -F_k + \left(d_{k-1} \left(\frac{y_{k-1}^\top}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{s_{k-1}^\top}{d_{k-1}^\top y_{k-1}} \right) \right) F_k \\
&= -F_k + d_{k-1} \left(\frac{y_{k-1}^\top F_k}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{s_{k-1}^\top F_k}{d_{k-1}^\top y_{k-1}} \right) \\
&= -F_k + d_{k-1} \left(\frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right) \\
&= -F_k + d_{k-1} \beta_k(\tau_{k-1}) \\
&= -F_k + \beta_k^+(\tau_{k-1}) d_{k-1} \\
&= d_k.
\end{aligned}$$

Para prosseguir com a análise, seja \bar{H}_k^P dado por

$$\bar{H}_k^P := \frac{1}{2} [H_k^P + (H_k^P)^\top] = I_n - \frac{d_{k-1} b_{k-1}^\top + b_{k-1} d_{k-1}^\top}{2d_{k-1}^\top y_{k-1}}.$$

Note que \bar{H}_k^P é uma matriz simétrica, além disso

$$-\frac{d_{k-1} b_{k-1}^\top + b_{k-1} d_{k-1}^\top}{2d_{k-1}^\top y_{k-1}} = \bar{H}_k^P - I_n = \bar{H}_k^P + (-I_n),$$

devido a isso, pela Proposição 1.4 segue que a matriz

$$-\frac{d_{k-1} b_{k-1}^\top + b_{k-1} d_{k-1}^\top}{2d_{k-1}^\top y_{k-1}},$$

também é simétrica. Graças a isso, podemos aplicar o Teorema Espectral 1.1, e assim afirmar que a matriz acima é ortogonalmente diagonalizável, isto é, diagonalizável. Observe que essa matriz é da forma

$$-\frac{uv^\top + vu^\top}{2d_{k-1}^\top y_{k-1}},$$

com $u = d_{k-1}$ e $v = b_{k-1}$ por isso, pelo Corolário 1.1 sabemos que o seu posto é, no máximo, 2. Como ela é semelhante a uma matriz diagonal (pois é diagonalizável) onde cada elemento não nulo da matriz diagonal são os autovalores dessa matriz (de acordo com a Proposição 1.8) e pelo fato de matrizes semelhantes terem o mesmo posto (de acordo com a Proposição 1.9) concluímos que essa matriz possui os autovalores $\lambda_1, \lambda_2 \in \mathbb{R}$ e 0. Devido a sua forma $(uv^\top + vu^\top)$, sabemos ainda, pela Proposição 1.12, que

$$\lambda_1 = -\frac{d_{k-1}^\top b_{k-1} + \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}} \quad \text{e} \quad \lambda_2 = -\frac{d_{k-1}^\top b_{k-1} - \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}}.$$

Por isso, aplicando a Proposição 1.10 com $\beta = \gamma = 1$ sabemos que \bar{H}_k^P possui, no máximo, três autovalores distintos $1, 1 + \lambda_1$ e $1 + \lambda_2$, isto é,

$$1, 1 - \frac{d_{k-1}^\top b_{k-1} - \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}} \quad \text{e} \quad 1 - \frac{d_{k-1}^\top b_{k-1} + \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}}.$$

Além disso, para quaisquer $u, v \in \mathbb{R}^n$, temos que $u^\top v - \|u\| \|v\| < u^\top v + \|u\| \|v\|$. Fazendo, $u = d_{k-1}$ e $v = b_{k-1}$. Se $d_{k-1}^\top y_{k-1} < 0$, então $-\frac{1}{2} \left(\frac{1}{d_{k-1}^\top y_{k-1}} \right) > 0$, logo

$$1 - \frac{d_{k-1}^\top b_{k-1} - \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}} < 1 - \frac{d_{k-1}^\top b_{k-1} + \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}},$$

além disso,

$$\begin{aligned}
1 - \frac{d_{k-1}^\top b_{k-1} - \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}} &= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1} - \|d_{k-1}\| \|b_{k-1}\|}{d_{k-1}^\top y_{k-1}} \right) \\
&= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} - \frac{\|d_{k-1}\| \|b_{k-1}\|}{d_{k-1}^\top y_{k-1}} \right) \\
&= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} + \frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \right).
\end{aligned}$$

Analogamente, se $d_{k-1}^\top y_{k-1} > 0$, então $-\frac{1}{2} \left(\frac{1}{d_{k-1}^\top y_{k-1}} \right) < 0$, logo

$$1 - \frac{d_{k-1}^\top b_{k-1} + \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}} < 1 - \frac{d_{k-1}^\top b_{k-1} - \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}},$$

assim,

$$\begin{aligned}
1 - \frac{d_{k-1}^\top b_{k-1} + \|d_{k-1}\| \|b_{k-1}\|}{2d_{k-1}^\top y_{k-1}} &= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1} + \|d_{k-1}\| \|b_{k-1}\|}{d_{k-1}^\top y_{k-1}} \right) \\
&= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} + \frac{\|d_{k-1}\| \|b_{k-1}\|}{d_{k-1}^\top y_{k-1}} \right) \\
&= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} + \frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \right).
\end{aligned}$$

Graças aos fatos citados anteriormente, não é difícil ver que o autovalor minimal de \bar{H}_k^P é

$$\lambda_{\min} = \min \left\{ 1, 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} + \frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \right) \right\}. \quad (12)$$

Com as definições de $p_{k-1} = \frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2}$ e $r_{k-1} = \tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}}$, podemos reescrever (12) como

$$\lambda_{\min} = \min\{1, \lambda\}, \quad (13)$$

onde

$$\lambda = \frac{1}{2} \left\{ p_{k-1} + r_{k-1} - \sqrt{p_{k-1}^2 + (2r_{k-1} - 3)p_{k-1} + (r_{k-1}^2 - 4r_{k-1} + 3)} \right\}.$$

Note que

$$\begin{aligned}
\frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} &= \frac{\|d_{k-1}\| \|b_{k-1}\|}{\sqrt{(d_{k-1}^\top y_{k-1})^2}} \\
&= \sqrt{\frac{\|d_{k-1}\|^2 \|b_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2}} \\
&= \sqrt{\frac{\|d_{k-1}\|^2 b_{k-1}^\top b_{k-1}}{(d_{k-1}^\top y_{k-1})^2}} \\
&= \sqrt{\frac{\|d_{k-1}\|^2 (y_{k-1}^\top - c_{k-1} s_{k-1}^\top)(y_{k-1} - c_{k-1} s_{k-1})}{(d_{k-1}^\top y_{k-1})^2}} \\
&= \sqrt{\frac{\|d_{k-1}\|^2 (\|y_{k-1}\|^2 - 2c_{k-1} s_{k-1}^\top y_{k-1} + c_{k-1}^2 \|s_{k-1}\|^2)}{(d_{k-1}^\top y_{k-1})^2}} \\
&= \sqrt{\frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2 - 2c_{k-1} \|d_{k-1}\|^2 s_{k-1}^\top y_{k-1} + c_{k-1}^2 \|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2}} \\
&= \sqrt{\frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} - 2c_{k-1} \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} + c_{k-1}^2 \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2}}.
\end{aligned}$$

Antes de prosseguirmos com os cálculos, precisamos estabelecer as seguintes igualdades,

$$\begin{aligned}
c_{k-1} \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} &= \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} \\
&= \tau_{k-1} \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} \\
&= \tau_{k-1} \frac{\|s_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(s_{k-1}^\top y_{k-1})^2} + \frac{\|y_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{\|s_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(s_{k-1}^\top y_{k-1})^2} \\
&= \tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + p_{k-1} - 1 \\
&= r_{k-1} + p_{k-1} - 1,
\end{aligned}$$

$$\begin{aligned}
c_{k-1}^2 \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} &= \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right)^2 \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} \\
&= \left(\tau_{k-1}^2 + \frac{\|y_{k-1}\|^4}{(s_{k-1}^\top y_{k-1})^2} + \frac{(s_{k-1}^\top y_{k-1})^2}{\|s_{k-1}\|^4} + 2\tau_{k-1} \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - 2\tau_{k-1} \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} - 2 \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} \\
&= \tau_{k-1}^2 \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} + \frac{\|y_{k-1}\|^4}{(s_{k-1}^\top y_{k-1})^2} \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} + \frac{(s_{k-1}^\top y_{k-1})^2}{\|s_{k-1}\|^4} \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} + 2\tau_{k-1} \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} - 2\tau_{k-1} \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} - 2 \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} \\
&= \tau_{k-1}^2 \frac{\|s_{k-1}\|^2 \|s_{k-1}\|^2}{(s_{k-1}^\top y_{k-1})^2} + \frac{\|y_{k-1}\|^4}{(d_{k-1}^\top y_{k-1})^2} \frac{\|d_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} + \frac{(s_{k-1}^\top y_{k-1})^2}{\|s_{k-1}\|^4} \frac{\|s_{k-1}\|^2 \|s_{k-1}\|^2}{(s_{k-1}^\top y_{k-1})^2} + 2\tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} - 2\tau_{k-1} \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{\|s_{k-1}\|^2 \|s_{k-1}\|^2}{(s_{k-1}^\top y_{k-1})^2} - 2 \frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} \\
&= \tau_{k-1}^2 \frac{\|s_{k-1}\|^4}{(s_{k-1}^\top y_{k-1})^2} + \frac{\|d_{k-1}\|^4 \|y_{k-1}\|^4}{(d_{k-1}^\top y_{k-1})^4} + 1 + 2r_{k-1} p_{k-1} - 2\tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - 2p_{k-1} \\
&= r_{k-1}^2 + p_{k-1}^2 + 1 + 2r_{k-1} p_{k-1} - 2r_{k-1} - 2p_{k-1}.
\end{aligned}$$

Com isso, temos

$$\begin{aligned} \frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} &= \sqrt{\frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} - 2c_{k-1} \frac{\|d_{k-1}\|^2 s_{k-1}^\top y_{k-1}}{(d_{k-1}^\top y_{k-1})^2} + c_{k-1}^2 \frac{\|s_{k-1}\|^2 \|d_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2}} \\ &= \sqrt{p_{k-1} - 2(r_{k-1} + p_{k-1} - 1) + (r_{k-1}^2 + p_{k-1}^2 + 1 + 2r_{k-1}p_{k-1} - 2r_{k-1} - 2p_{k-1})} \\ &= \sqrt{p_{k-1}^2 + (2r_{k-1} - 3)p_{k-1} + (r_{k-1}^2 - 4r_{k-1} + 3)}. \end{aligned}$$

Prosseguindo, repare que

$$\begin{aligned} c_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} &= \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \\ &= \tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \\ &= r_{k-1} + \frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(s_{k-1}^\top y_{k-1})^2} - 1 \\ &= r_{k-1} + p_{k-1} + 1. \end{aligned}$$

Logo, $p_{k-1} + r_{k-1} = c_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + 1$. Por isso,

$$\begin{aligned} 1 - \frac{1}{2} \frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} &= 1 - \frac{1}{2} \frac{d_{k-1}^\top (y_{k-1} - c_{k-1} s_{k-1})}{d_{k-1}^\top y_{k-1}} \\ &= 1 - \frac{1}{2} \frac{d_{k-1}^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} + \frac{1}{2} c_{k-1} \frac{d_{k-1}^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \\ &= 1 - \frac{1}{2} + \frac{1}{2} c_{k-1} \frac{s_{k-1}^\top s_{k-1}}{s_{k-1}^\top y_{k-1}} \\ &= \frac{1}{2} + \frac{1}{2} c_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \\ &= \frac{1}{2} \left(c_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + 1 \right) \\ &= \frac{1}{2} (p_{k-1} + r_{k-1}). \end{aligned}$$

Consequentemente,

$$\begin{aligned} \lambda &= 1 - \frac{1}{2} \left(\frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} + \frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \right) \\ &= \left(1 - \frac{1}{2} \frac{d_{k-1}^\top b_{k-1}}{d_{k-1}^\top y_{k-1}} \right) - \frac{1}{2} \frac{\|d_{k-1}\| \|b_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \\ &= \frac{1}{2} (p_{k-1} + r_{k-1}) - \frac{1}{2} \sqrt{p_{k-1}^2 + (2r_{k-1} - 3)p_{k-1} + (r_{k-1}^2 - 4r_{k-1} + 3)} \\ &= \frac{1}{2} \left\{ p_{k-1} + r_{k-1} - \sqrt{p_{k-1}^2 + (2r_{k-1} - 3)p_{k-1} + (r_{k-1}^2 - 4r_{k-1} + 3)} \right\}. \end{aligned}$$

Chamando $x = p_{k-1}$ e $r = r_{k-1}$, considere a função,

$$f(x) = \lambda = \frac{1}{2} \left\{ x + r - \sqrt{x^2 + (2r - 3)x + (r^2 - 4r + 3)} \right\},$$

Para saber qual é o domínio da função acima, basta notar que da desigualdade de Cauchy-Schwarz temos

$$|d_{k-1}^\top y_{k-1}| \leq \|d_{k-1}\| \|y_{k-1}\|,$$

como ambos os membros da desigualdade acima são não-negativos, podemos quadrar a mesma sem inverter o sentido da inequação e também não alterar o seu conjunto solução consequentemente,

$$|d_{k-1}^\top y_{k-1}|^2 = (d_{k-1}^\top y_{k-1})^2 \leq \|d_{k-1}\|^2 \|y_{k-1}\|^2,$$

o que implica,

$$\frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} \geq 1.$$

Observe, que a fração acima é $p_{k-1} = x$, portanto o domínio de f é o intervalo $[1, +\infty)$. Retomando, observe a equação abaixo

$$f(x) = \frac{1}{2} \left\{ x + r - \sqrt{x^2 + (2r - 3)x + (r^2 - 4r + 3)} \right\},$$

do polinômio $x^2 + (2r - 3)x + (r^2 - 4r + 3)$ observamos que para cada x do domínio, o polinômio assumirá valores não-negativos, desde que

$$(2r - 3)^2 - 4(r^2 - 4r + 3) \leq 0 \Leftrightarrow (4r^2 - 12r + 9) - (4r^2 - 16r + 12) \leq 0 \Leftrightarrow 4r - 3 \leq 0 \Leftrightarrow r \leq \frac{3}{4},$$

e a primeira coordenada de seu vértice é,

$$x_v = -\frac{2r-3}{2} = \frac{3}{2} - r.$$

De,

$$f(x) = \frac{1}{2} \left\{ x + r - \sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)} \right\},$$

podemos encontrar a sua derivada,

$$f'(x) = \frac{1}{2} - \frac{1}{2} \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}}.$$

Feito estas observações, queremos mostrar que f tem mínimo global em $x = 1$, desde que $r \leq \frac{3}{4}$ para isto, vamos considerar dois casos, se $r = \frac{3}{4}$ é importante destacar que

$$f'(x) = 0 \Leftrightarrow \frac{1}{2} = \frac{1}{2} \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}} \Leftrightarrow 1 = \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}}.$$

Vamos definir,

$$g(x) = \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}},$$

para mostrar que f é constante, basta verificar que $g(x) = 1$ para todo x do domínio. Para este fim,

$$g'(x) = \frac{\frac{3}{4} - r}{x^2 + (2r-3)x + (r^2 - 4r + 3)},$$

como $r = \frac{3}{4}$, observamos que $g'(x) = 0$, além disso

$$g(x) = \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}} = \frac{x - \frac{3}{4}}{\sqrt{x^2 - \frac{3}{2}x + \frac{9}{16}}},$$

ou seja,

$$g(1) = \frac{1 - \frac{3}{4}}{\sqrt{1 - \frac{3}{2} + \frac{9}{16}}} = \frac{\frac{1}{4}}{\sqrt{\frac{1}{16}}} = 1.$$

Com isso, f é constante quando $r = \frac{3}{4}$, assim qualquer valor de x é o valor mínimo de f . Caso $r < \frac{3}{4}$, vamos mostrar que f é crescente, para isso,

$$f'(x) > 0 \Leftrightarrow 1 > \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}}.$$

Para mostrar que f é crescente, basta verificar que $g(x) < 1$ para todo x do domínio. Para este fim, lembre-se que

$$g'(x) = \frac{\frac{3}{4} - r}{x^2 + (2r-3)x + (r^2 - 4r + 3)},$$

como $r < \frac{3}{4}$, tanto o numerador e o denominador são estritamente positivos, assim $g'(x) > 0$, ou seja g é estritamente crescente em seu domínio, além disso,

$$\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} \frac{x + r - \frac{3}{2}}{\sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)}} = \lim_{x \rightarrow \infty} \frac{1 + \frac{r}{x} - \frac{3}{2x}}{\sqrt{1 + \frac{2r-3}{x} + \frac{(r^2-4r+3)}{x^2}}} = 1.$$

Com essas duas informações, observe que não existe nenhum x do domínio tal que $g(x) > 1$, pois caso contrário como g é contínua o teorema do valor intermediário garantiria a existência de um $c > x$ tal que $g(x) > g(c) > 1$, ou seja g seria decrescente no intervalo (x, c) o que contradiz o fato de g ser estritamente crescente.

Por outro lado, se tivéssemos $g(c) = 1$, considerando o caso anterior, caso esse c fosse único existiria um intervalo do qual g decresceria, caso c não fosse único o teorema de Rolle garantiria a existência de um d tal que $g'(d) = 0$, ambos os casos não podem acontecer e portanto, $g(x) < 1$ para todo x do domínio.

Donde,

$$f(1) = \frac{1}{2}(1 + r - |r - 1|) = \frac{1}{2}(1 + r - (1 - r)) = r.$$

Portanto, $f(x) \geq r$ caso $r \leq \frac{3}{4}$. Por outro lado, se $r > \frac{3}{4}$, vamos lembrar que

$$f(x) = \frac{1}{2} \left\{ x + r - \sqrt{x^2 + (2r-3)x + (r^2 - 4r + 3)} \right\},$$

vamos verificar as raízes do polinômio $x^2 + (2r-3)x + (r^2 - 4r + 3)$ que são,

$$x_1 = \frac{3 - 2r - \sqrt{4r - 3}}{2} \quad \text{e} \quad x_2 = \frac{3 - 2r + \sqrt{4r - 3}}{2}.$$

Evidentemente, $x_2 > x_1$ e ao resolver a inequação $x_2 > 1$, observamos que

$$\frac{3 - 2r + \sqrt{4r - 3}}{2} > 1 \Leftrightarrow \sqrt{4r - 3} > 2r - 1 \Leftrightarrow 4r - 3 > 4r^2 - 4r + 1 \Leftrightarrow 4(r - 1)^2 < 0,$$

o que não pode ocorrer, analogamente ao resolver a equação $x_2 = 1$, obtemos $r = 1$, logo $x_1 = 0$, ou seja f é definida para todo $x \geq 1$.

Por cálculos análogos, podemos afirmar que f é decrescente se $r > \frac{3}{4}$ com $r \neq 1$, da seguinte maneira. Primeiro verificamos para quais elementos do domínio vale $f'(x) < 0$ mas este problema é equivalente a mostrar que $g(x) > 1, \forall x \geq 1$, em que

$$g'(x) = \frac{\frac{3}{4} - r}{x^2 + (2r - 3)x + (r^2 - 4r + 3)},$$

note que $g'(x) < 0$ pois, se $r > \frac{3}{4}$, então $\frac{3}{4} - r < 0$ e $x^2 + (2r - 3)x + (r^2 - 4r + 3) > 0, \forall x \geq 1$ conseqüentemente, g é estritamente decrescente. Sabemos ainda que $g(x) \rightarrow 1$ assim, usando a continuidade da g podemos concluir que $g(x) > 1$ para todo x do domínio.

Temos ainda que,

$$\begin{aligned} \lim_{x \rightarrow \infty} f(x) &= \lim_{x \rightarrow \infty} \frac{1}{2} \left\{ x + r - \sqrt{x^2 + (2r - 3)x + (r^2 - 4r + 3)} \right\} \\ &= \frac{1}{2} \lim_{x \rightarrow \infty} x \left(1 + \frac{r}{x} - \sqrt{1 + \frac{2r - 3}{x} + \frac{r^2 - 4r + 3}{x^2}} \right) \\ &= \frac{1}{2} \lim_{x \rightarrow \infty} \frac{1 + \frac{r}{x} - \sqrt{1 + \frac{2r - 3}{x} + \frac{r^2 - 4r + 3}{x^2}}}{\frac{1}{x}} \\ &= \frac{1}{2} \lim_{x \rightarrow \infty} \frac{-\frac{r}{x^2} - \frac{\frac{3 - 2r}{x^2} - \frac{2r^2 - 8r + 6}{x^3}}{2\sqrt{1 + \frac{2r - 3}{x} + \frac{r^2 - 4r + 3}{x^2}}}}{-\frac{1}{x^2}} \\ &= \frac{1}{2} \lim_{x \rightarrow \infty} \left(r + \frac{3 - 2r}{2\sqrt{1 + \frac{2r - 3}{x} + \frac{r^2 - 4r + 3}{x^2}}} - \frac{2r^2 - 8r + 6}{2x\sqrt{1 + \frac{2r - 3}{x} + \frac{r^2 - 4r + 3}{x^2}}} \right) \\ &= \frac{1}{2} \left(r + \left(\frac{3}{2} - r \right) \right) \\ &= \frac{3}{4}. \end{aligned}$$

No caso em que $r = 1$ temos $f(x) = \frac{1}{2}(x + 1 - \sqrt{x^2 - x})$, logo $f(1) = 1$, além disso,

$$f'(x) = \frac{1}{2} - \frac{2x - 1}{4\sqrt{x^2 - x}} = \frac{2\sqrt{x^2 - x} - 2x + 1}{4\sqrt{x^2 - x}},$$

quando $x > 1$ notamos que $\sqrt{x^2 - x} > 0$, assim o denominador de f' é positivo, já no numerador observe que se,

$$0 < 2\sqrt{x^2 - x} < 2x + 1 \Leftrightarrow 4(x^2 - x) < 4x^2 + 4x + 1 \Leftrightarrow -1 < 8x \Leftrightarrow -\frac{1}{8} < x,$$

ou seja, f é estritamente decrescente em $(1, +\infty)$ e como $f \rightarrow \frac{3}{4}$, usando a continuidade de f concluímos que $f(x) > \frac{3}{4}$ para todo x do domínio quando $r = 1$.

Assim, se $r_{k-1} = r \leq \frac{3}{4}$, f tem seu mínimo em $x = p_{k-1} = 1$, ou seja, $f(x) = \lambda \geq r_{k-1}$. Por outro lado, acabamos de ver que se $r_{k-1} > \frac{3}{4}$, vamos ter que $f(x) = \lambda > \frac{3}{4}$. Juntando ambas as informações concluímos que,

$$\lambda \geq \min \left\{ r_{k-1}, \frac{3}{4} \right\},$$

por outro lado, sabemos que

$$\lambda_{\min} = \min\{1, \lambda\} \geq \min \left\{ 1, \frac{3}{4}, r_{k-1} \right\} = \min \left\{ r_{k-1}, \frac{3}{4} \right\}.$$

Observe que

$$F_k^\top H_k^P F_k = F_k^\top \left(I - \frac{d_{k-1} b_{k-1}^\top}{d_{k-1}^\top y_{k-1}} \right) F_k = \|F_k\|^2 - \frac{F_k^\top d_{k-1} b_{k-1}^\top F_k}{d_{k-1}^\top y_{k-1}},$$

conseqüentemente,

$$\begin{aligned} F_k^\top \bar{H}_k^P F_k &= F_k^\top \left(I - \frac{d_{k-1} b_{k-1}^\top + b_{k-1} d_{k-1}^\top}{2d_{k-1}^\top y_{k-1}} \right) F_k \\ &= \|F_k\|^2 - \frac{F_k^\top d_{k-1} b_{k-1}^\top F_k}{2d_{k-1}^\top y_{k-1}} - \frac{F_k^\top b_{k-1} d_{k-1}^\top F_k}{2d_{k-1}^\top y_{k-1}} \\ &= \|F_k\|^2 - \frac{F_k^\top d_{k-1} b_{k-1}^\top F_k}{d_{k-1}^\top y_{k-1}} \quad (\text{pela Proposição 1.13}) \\ &= F_k^\top H_k^P F_k. \end{aligned}$$

Ou seja, podemos usar o Teorema 1.2 da seguinte forma

$$F_k^\top d_k = -F_k^\top H_k^P F_k = -F_k^\top \bar{H}_k^P F_k \leq -\lambda_{\min} \|F_k\|^2 \leq -\min \left\{ r_{k-1}, \frac{3}{4} \right\} \|F_k\|^2.$$

E assim, provamos a primeira parte do primeiro caso deste lema. Além disso, como $\tau_{k-1} = \theta \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + (1-\theta) \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2}$, provaremos que existe uma constante $c_1 \in (0, 1)$, tal que $F_k^\top d_k \leq -c_1 \|F_k\|^2$, Se $\tau_{k-1} = \tau_{k-1}^A = \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}}$ então

$$r_{k-1} = \tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} = \frac{\|s_{k-1}\|^2 \|y_{k-1}\|^2}{(s_{k-1}^\top y_{k-1})^2} = \frac{\|d_{k-1}\|^2 \|y_{k-1}\|^2}{(d_{k-1}^\top y_{k-1})^2} = p_{k-1}.$$

Pelo fato de $\lambda_{\min} = \min\{1, \lambda\}$ e $p_{k-1} \geq 1$ podemos concluir que

$$\begin{aligned} \lambda_{\min} &= \min\{1, \lambda\} \\ &= \min \left\{ 1, \frac{1}{2} \left\{ p_{k-1} + r_{k-1} - \sqrt{p_{k-1}^2 + (2r_{k-1} - 3)p_{k-1} + (r_{k-1}^2 - 4r_{k-1} + 3)} \right\} \right\} \\ &= \min \left\{ 1, p_{k-1} - \frac{1}{2} \sqrt{4p_{k-1}^2 - 7p_{k-1} + 3} \right\} \\ &= \min \left\{ 1, p_{k-1} - \sqrt{p_{k-1}^2 - \frac{7}{4}p_{k-1} + \frac{3}{4}} \right\} \\ &> \min \left\{ 1, p_{k-1} - \sqrt{p_{k-1}^2 - \frac{7}{4}p_{k-1} + \frac{49}{64}} \right\}. \end{aligned}$$

Observe que

$$p_{k-1} - \sqrt{p_{k-1}^2 - \frac{7}{4}p_{k-1} + \frac{49}{64}} = p_{k-1} - \sqrt{\left(p_{k-1} - \frac{7}{8}\right)^2} = p_{k-1} - \left|p_{k-1} - \frac{7}{8}\right| = p_{k-1} - \left(p_{k-1} - \frac{7}{8}\right) = \frac{7}{8} < 1,$$

portanto, neste caso vale

$$\lambda_{\min} > \min \left\{ 1, p_{k-1} - \sqrt{p_{k-1}^2 - \frac{7}{4}p_{k-1} + \frac{49}{64}} \right\} = \frac{7}{8},$$

assim,

$$F_k^\top d_k \leq -\min \left\{ r_{k-1}, \frac{3}{4} \right\} \|F_k\|^2 = -\frac{7}{8} \|F_k\|^2.$$

Caso $\tau_{k-1} = \tau_{k-1}^B = \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2}$, observamos que $r_{k-1} = \tau_{k-1} \frac{\|s_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} = 1$, evidentemente teremos que $c_1 = \frac{3}{4}$.

Observe que $\beta_k(\tau_{k-1})$ é uma função afim,

$$\beta_k(\tau_{k-1}) = \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} = a\tau_{k-1} + b,$$

em que $a, b \in \mathbb{R}$. Logo, $F_k^\top d_k$ é também função afim na variável τ_{k-1} pois,

$$F_k^\top d_k = F_k^\top (-F_k + \beta_k^+(\tau_{k-1})d_{k-1}) = F_k^\top (-F_k + (a\tau_{k-1} + b)d_{k-1}) = -\|F_k\|^2 + (aF_k^\top d_{k-1})\tau_{k-1} + bF_k^\top d_{k-1} = \alpha\tau_{k-1} + \beta.$$

Note que $\alpha = aF_k^\top d_{k-1}$ e $\beta = bF_k^\top d_{k-1} + \|F_k\|^2$. Graficamente, $F_k^\top d_k$ localiza-se no hipográfico de um segmento de reta que passa pelos pontos $-\frac{7}{8}\|F_k\|^2$ e $-\frac{3}{4}\|F_k\|^2$ ou seja, para $0 < \theta < 1$, concluímos que $F_k^\top d_k \leq -\frac{3}{4}\|F_k\|^2$, isso completa o primeiro caso.

Se $\beta_k(\tau_{k-1}) < \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2}$, então $\beta_k^+(\tau_{k-1}) = \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2}$, conseqüentemente

$$\begin{aligned} F_k^\top d_k &= F_k^\top (-F_k + \beta_k(\tau_{k-1})d_{k-1}) \\ &= -\|F_k\|^2 + \beta_k(\tau_{k-1})F_k^\top d_{k-1} \\ &= -\|F_k\|^2 + \left(\eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2} \right) F_k^\top d_{k-1} \\ &= -\|F_k\|^2 + \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2} \\ &\leq -\|F_k\|^2 + \eta \frac{\|F_k\|^2 \|d_{k-1}\|^2}{\|d_{k-1}\|^2} \\ &= -\|F_k\|^2 + \eta \|F_k\|^2 \\ &= -(1-\eta)\|F_k\|^2. \end{aligned}$$

Pelo fato de $0 \leq \eta < 1$, temos que $-1 < -\eta \leq 0$, logo $0 < 1 - \eta \leq 1$, isso completa a demonstração. \square

Quando $\eta = 0$ temos que $F_k^\top d_k = -\|F_k\|^2$. Apesar desse valor de η não satisfazer o Lema 3.1 os valores de η restantes os satisfazem. Com isso, podemos concluir que apesar do Algoritmo 1 mostrar que $\eta \in [0, 1)$, tomar $\eta = 0$ é desaconselhável, visto que a análise de convergência do método apresentada na próxima seção utilizará esse Lema, ou seja, ao tomar $\eta = 0$ nada podemos afirmar sobre a convergência da sequência $\{\|F_k\|\}$.

4 Análise de Convergência

A convergência global do Algoritmo 1 foi verificada assumindo que a função F satisfaz as seguintes hipóteses.

Hipótese 1: A função F é Lipschitz contínua em Ω , isto é, existe uma constante $L > 0$ de modo que

$$\|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, y \in \Omega.$$

Hipótese 2: A função $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é continuamente diferenciável e monótona, assim por consequência desta última característica: $(F(x) - F(y))^\top (x - y) \geq 0$, $\forall x, y \in \mathbb{R}^n$.

Hipótese 3: O conjunto solução de $F(x) = 0$, $x \in \Omega$ que será denotado por Φ é não-vazio. Para estabelecermos a convergência do Algoritmo, primeiro serão demonstrados os seguintes lemas.

Lema 4.1. *A busca linear (10) é bem definida.*

Demonstração. Se o algoritmo termina em alguma iterada, digamos k , então $\|F_k\| = 0$ (ou $F(z_k) = 0$), logo x_k (ou z_k) é solução. Assuma que $F_k \neq 0$ para todo k , e consequentemente $d_k \neq 0$ pelo Lema 3.1. Agora mostraremos que a busca linear feita no passo 2 do algoritmo sempre termina em um número finito de passos. Se $t_k \neq \xi$, sabemos que $t'_k = \rho^{-1}t_k$ não satisfará a busca, isto é

$$-F(z'_k)^\top d_k < \sigma t'_k \|d_k\|^2,$$

em que $z'_k = x_k + t'_k d_k$.

Pelo Lema 3.1 e pelo fato de F ser Lipschitz contínua em Ω temos

$$\begin{aligned} c\|F_k\|^2 &\leq -F_k^\top d_k \\ &= (F(z'_k) - F(x_k))^\top d_k - F(z'_k)^\top d_k \\ &< (F(z'_k) - F(x_k))^\top d_k + \sigma t'_k \|d_k\|^2 \\ &\leq |(F(z'_k) - F(x_k))^\top d_k| + \sigma t'_k \|d_k\|^2 \\ &\leq \|F(z'_k) - F(x_k)\| \|d_k\| + \sigma t'_k \|d_k\|^2 \\ &\leq L\|z'_k - x_k\| \|d_k\| + \sigma t'_k \|d_k\|^2 \\ &= L\|(x_k + t'_k d_k) - x_k\| \|d_k\| + \sigma t'_k \|d_k\|^2 \\ &= L t'_k \|d_k\| \|d_k\| + \sigma t'_k \|d_k\|^2 \\ &= L t'_k \|d_k\|^2 + \sigma t'_k \|d_k\|^2 \\ &= t'_k (L + \sigma) \|d_k\|^2 \\ &= \rho^{-1} t_k (L + \sigma) \|d_k\|^2. \end{aligned}$$

Ou seja,

$$c\|F_k\|^2 = \rho^{-1} t_k (L + \sigma) \|d_k\|^2.$$

O que acarreta,

$$t_k = \frac{c\rho}{L + \sigma} \frac{\|F_k\|^2}{\|d_k\|^2}.$$

Portanto, podemos concluir que

$$t_k \geq \min \left\{ \xi, \frac{c\rho}{L + \sigma} \frac{\|F_k\|^2}{\|d_k\|^2} \right\},$$

e assim segue que a busca linear (10) é bem definida. \square

Lema 4.2. *Suponha que F é Lipschitz contínua em Ω , continuamente diferenciável e monótona, além disso $\Phi \neq \emptyset$ e a sequência $\{x_k\}$ gerada pelo Algoritmo 1. Existe uma constante estritamente positiva M , tal que*

$$\|F(x_k)\| \leq M. \tag{14}$$

Demonstração. Para qualquer $\bar{x} \in \Phi$, usando a não-expansividade do operador projeção, isto é, $\|P_\Omega[x] - P_\Omega[y]\| \leq \|x - y\|$ para quaisquer $x, y \in \mathbb{R}^n$, temos

$$\begin{aligned} \|x_{k+1} - \bar{x}\|^2 &= \|P_\Omega[x_k - \alpha F(z_k)] - \bar{x}\|^2 \\ &= \|P_\Omega[x_k - \alpha F(z_k)] - P_\Omega[\bar{x}]\|^2 \\ &\leq \|x_k - \alpha F(z_k) - \bar{x}\|^2 \\ &= \langle (x_k - \bar{x}) - \alpha F(z_k), (x_k - \bar{x}) - \alpha F(z_k) \rangle \\ &= \langle (x_k - \bar{x}), (x_k - \bar{x}) - \alpha F(z_k) \rangle - \langle \alpha F(z_k), (x_k - \bar{x}) - \alpha F(z_k) \rangle \\ &= \langle (x_k - \bar{x}), (x_k - \bar{x}) \rangle - \langle (x_k - \bar{x}), \alpha F(z_k) \rangle - \langle \alpha F(z_k), (x_k - \bar{x}) \rangle + \langle \alpha F(z_k), \alpha F(z_k) \rangle \\ &= \langle (x_k - \bar{x}), (x_k - \bar{x}) \rangle - 2\langle \alpha F(z_k), (x_k - \bar{x}) \rangle + \langle \alpha F(z_k), \alpha F(z_k) \rangle \\ &= \|x_k - \bar{x}\|^2 - 2\alpha \langle F(z_k), (x_k - \bar{x}) \rangle + \alpha^2 \|F(z_k)\|^2. \end{aligned}$$

Pela monotonicidade da função F temos

$$(F(z_k) - F(\bar{x}))^\top (z_k - \bar{x}) \geq 0,$$

o que acarreta

$$F(z_k)^\top (z_k - \bar{x}) - F(\bar{x})^\top (z_k - \bar{x}) \geq 0,$$

mas como $\bar{x} \in \Phi$ repare que $F(\bar{x})^\top = 0^\top$, consequentemente

$$F(z_k)^\top (z_k - \bar{x}) \geq 0,$$

da igualdade $z_k - \bar{x} = (x_k - \bar{x}) - (x_k - z_k)$ podemos reescrever a desigualdade acima como

$$F(z_k)^\top ((x_k - \bar{x}) - (x_k - z_k)) \geq 0,$$

logo

$$F(z_k)^\top(x_k - \bar{x}) \geq F(z_k)^\top(x_k - z_k).$$

Além disso,

$$\alpha_k = \frac{F(z_k)^\top(x_k - z_k)}{\|F(z_k)\|^2} = \frac{F(z_k)^\top(x_k - (x_k + t_k d_k))}{\|F(z_k)\|^2} = \frac{-t_k F(z_k)^\top d_k}{\|F(z_k)\|^2} \geq \frac{\sigma t_k^2 \|d_k\|^2}{\|F(z_k)\|^2} > 0,$$

portanto,

$$-2\alpha_k F(z_k)^\top(x_k - \bar{x}) \leq -2\alpha_k F(z_k)^\top(x_k - z_k).$$

Para retomarmos a demonstração, basta relembrarmos que,

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - 2\alpha_k F(z_k)^\top(x_k - \bar{x}) + \alpha_k^2 \|F(z_k)\|^2,$$

prossequindo com os cálculos usando a desigualdade obtida na observação anterior obtemos,

$$\begin{aligned} \|x_{k+1} - \bar{x}\|^2 &\leq \|x_k - \bar{x}\|^2 - 2\alpha_k F(z_k)^\top(x_k - z_k) + \alpha_k^2 \|F(z_k)\|^2 \\ &= \|x_k - \bar{x}\|^2 - 2 \left(\frac{F(z_k)^\top(x_k - z_k)}{\|F(z_k)\|^2} \right) F(z_k)^\top(x_k - z_k) + \left(\frac{F(z_k)^\top(x_k - z_k)}{\|F(z_k)\|^2} \right)^2 \|F(z_k)\|^2 \\ &= \|x_k - \bar{x}\|^2 - \frac{[F(z_k)^\top(x_k - z_k)]^2}{\|F(z_k)\|^2} \\ &\leq \|x_k - \bar{x}\|^2. \end{aligned} \tag{15}$$

Por cálculos análogos podemos obter,

$$0 \leq \|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 \leq \|x_{k-1} - \bar{x}\|^2 \leq \dots \leq \|x_0 - \bar{x}\|^2.$$

O que implica,

$$\|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\| \leq \|x_{k-1} - \bar{x}\| \leq \dots \leq \|x_0 - \bar{x}\|. \tag{16}$$

Pelo fato de F ser lipschitz contínua concluímos que,

$$\|F(x_k)\| = \|F(x_k) - F(\bar{x})\| \leq L\|x_k - \bar{x}\| \leq L\|x_0 - \bar{x}\|.$$

Tomando $M = L\|x_0 - \bar{x}\|$ podemos afirmar, de acordo com os argumentos supracitados, que encontramos uma constante estritamente estritamente positiva M , tal que $\|F(x_k)\| \leq M$. \square

Lema 4.3. *Suponha que F é lipschitz contínua em Ω , continuamente diferenciável e monótona, além disso $\Phi \neq \emptyset$ e as seqüências $\{x_k\}$ e $\{d_k\}$ geradas pelo Algoritmo 1. Nestas condições, podemos afirmar que $-F(z_k)$ é direção de descida da função $\frac{1}{2}\|x - \bar{x}\|^2$ no ponto x_k , de modo que $\bar{x} \in \Phi$.*

Demonstração. Primeiramente, o gradiente da função $f(x) = \frac{1}{2}\|x - \bar{x}\|^2$ no ponto x_k é $x_k - \bar{x}$, pois tomando $x = (x_1, x_2, \dots, x_n)$ e $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ observamos,

$$\frac{1}{2}\|x - \bar{x}\|^2 = \frac{1}{2} [(x_1 - \bar{x}_1)^2 + (x_2 - \bar{x}_2)^2 + \dots + (x_n - \bar{x}_n)^2],$$

o que implica,

$$\frac{\partial f}{\partial x_i} \left(\frac{1}{2}\|x - \bar{x}\|^2 \right) = x_i - \bar{x}_i, \quad i = 1, 2, \dots, n.$$

Consequentemente,

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) = (x_1 - \bar{x}_1, x_2 - \bar{x}_2, \dots, x_n - \bar{x}_n) = (x_1, x_2, \dots, x_n) - (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = x - \bar{x}.$$

Desta forma,

$$\nabla f(x_k)^\top F(z_k) = \langle x_k - \bar{x}, F(z_k) \rangle = \langle (x_k - z_k) + (z_k - \bar{x}), F(z_k) \rangle = \langle x_k - z_k, F(z_k) \rangle + \langle z_k - \bar{x}, F(z_k) \rangle.$$

Note que pela monotonicidade,

$$(F(z_k) - F(\bar{x}))^\top(z_k - \bar{x}) \geq 0,$$

mas como $(F(z_k) - F(\bar{x}))^\top(z_k - \bar{x}) = F(z_k)^\top(z_k - \bar{x}) = \langle z_k - \bar{x}, F(z_k) \rangle$ e devido ao fato de $F(\bar{x})^\top(z_k - \bar{x}) = \langle z_k - \bar{x}, F(\bar{x}) \rangle = 0$ podemos escrever

$$\langle z_k - \bar{x}, F(z_k) \rangle \geq \langle z_k - \bar{x}, F(\bar{x}) \rangle.$$

Assim,

$$\begin{aligned} \nabla f(x_k)^\top F(z_k) &= \langle x_k - z_k, F(z_k) \rangle + \langle z_k - \bar{x}, F(z_k) \rangle \\ &\geq \langle x_k - z_k, F(z_k) \rangle + \langle z_k - \bar{x}, F(\bar{x}) \rangle \\ &= \langle x_k - z_k, F(z_k) \rangle \\ &= \langle -t_k d_k, F(z_k) \rangle \\ &\geq \sigma t_k^2 \|d_k\|^2 \\ &= \sigma \|t_k d_k\|^2 \\ &= \sigma \|z_k - x_k\|^2 \\ &= \sigma \|x_k - z_k\|^2 \\ &> 0. \end{aligned} \tag{17}$$

Portanto, $\nabla f(x_k)^\top(-F(z_k)) < 0$ e consequentemente, mostramos que $-F(z_k)$ é direção de descida da função $f(x) = \frac{1}{2}\|x - \bar{x}\|^2$ no ponto x_k . \square

Lema 4.4. *Suponha que F é lipschitz contínua em Ω , continuamente diferenciável e monótona, além disso $\Phi \neq \emptyset$ e as seqüências $\{x_k\}$ e $\{d_k\}$ geradas pelo Algoritmo 1. Nestas condições, podemos afirmar*

- (1) *As seqüências $\{x_k\}$ e $\{z_k\}$ são limitadas;*
- (2) *$\lim_{k \rightarrow \infty} (x_k - z_k) = 0$, particularmente, $\lim_{k \rightarrow \infty} t_k \|d_k\| = 0$;*
- (3) *$\lim_{k \rightarrow \infty} (x_k - x_{k+1}) = 0$.*

Demonstração. (1) No Lema 4.3, vimos que

$$\|x_k - \bar{x}\| \leq \|x_{k-1} - \bar{x}\| \leq \dots \leq \|x_0 - \bar{x}\|,$$

como $\|x_k\| - \|\bar{x}\| \leq \|x_k - \bar{x}\|$ podemos escrever

$$\|x_k\| \leq \|x_0 - \bar{x}\| + \|\bar{x}\| = \|x_0 - \bar{x}\|.$$

Devido ao fato de existir um número real positivo, a saber, $\|x_0 - \bar{x}\|$ que limita superiormente $\|x_k\|$ concluímos que $\{x_k\}$ é limitada.

No Lema 4.3 vimos a seguinte desigualdade,

$$\langle F(z_k), x_k - z_k \rangle \geq \sigma \|x_k - z_k\|^2.$$

Por outro lado,

$$\langle F(z_k), x_k - z_k \rangle = \langle F(z_k) - F(x_k), x_k - z_k \rangle + \langle F(x_k), x_k - z_k \rangle,$$

pela monotonicidade de F notamos que $\langle F(z_k) - F(x_k), z_k - x_k \rangle \geq 0$ o que acarreta

$$-\langle F(z_k) - F(x_k), z_k - x_k \rangle = \langle F(z_k) - F(x_k), -(z_k - x_k) \rangle = \langle F(z_k) - F(x_k), x_k - z_k \rangle \leq 0,$$

assim,

$$\langle F(z_k), x_k - z_k \rangle \leq \langle F(x_k), x_k - z_k \rangle \leq |\langle F(x_k), x_k - z_k \rangle| \leq \|F(x_k)\| \|x_k - z_k\| \leq M \|x_k - z_k\|.$$

Repare que na última desigualdade o Lema 4.2 foi utilizado. Combinando as desigualdades $\sigma \|x_k - z_k\|^2 \leq \langle F(z_k), x_k - z_k \rangle$ e $\langle F(z_k), x_k - z_k \rangle \leq M \|x_k - z_k\|$ temos,

$$\sigma \|x_k - z_k\|^2 \leq M \|x_k - z_k\|,$$

com a informação acima, podemos deduzir que,

$$\|z_k\| - \|x_k\| \leq \|z_k - x_k\| = \|x_k - z_k\| \leq \frac{M}{\sigma},$$

desse modo, podemos obter a desigualdade,

$$\|z_k\| \leq \frac{M}{\sigma} + \|x_k\|.$$

Portanto, a seqüência $\{z_k\}$ é limitada pelo fato da seqüência $\{x_k\}$ também ser limitada. □

Demonstração. (2) No lema 4.2 mostramos que a desigualdade (15) é válida,

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \frac{[F(z_k)^\top (x_k - z_k)]^2}{\|F(z_k)\|^2},$$

e como

$$\langle F(z_k), x_k - z_k \rangle = F(z_k)^\top (x_k - z_k) \geq \sigma \|x_k - z_k\|^2.$$

É possível escrever que,

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \frac{\sigma^2 \|x_k - z_k\|^4}{\|F(z_k)\|^2},$$

o que acarreta,

$$\|x_k - z_k\|^4 \leq \frac{\|F(z_k)\|^2}{\sigma^2} (\|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2).$$

Pelo fato da função $F(x)$ ser contínua e a seqüência $\{z_k\}$ ser limitada, então a seqüência $\{\|F(z_k)\|\}$ também é limitada. Portanto, existe uma constante positiva $M_1 > 0$ tal que $\|F(z_k)\| \leq M_1$, desta forma

$$\|x_k - z_k\|^4 \leq \frac{\|F(z_k)\|^2}{\sigma^2} (\|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2) \leq \frac{M_1^2}{\sigma^2} (\|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2).$$

Definindo $s_k = \sum_{i=0}^k \frac{M_1^2}{\sigma^2} (\|x_i - \bar{x}\|^2 - \|x_{i+1} - \bar{x}\|^2)$ observamos que,

$$s_k = \sum_{i=0}^k \frac{M_1^2}{\sigma^2} (\|x_i - \bar{x}\|^2 - \|x_{i+1} - \bar{x}\|^2) = \frac{M_1^2}{\sigma^2} (\|x_0 - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2) \leq \frac{M_1^2}{\sigma^2} (\|x_0 - \bar{x}\|^2 + \|x_0 - \bar{x}\|^2) = 2 \frac{M_1^2}{\sigma^2} \|x_0 - \bar{x}\|^2.$$

Observe que a seqüência $\{s_k\}$ é limitada superiormente por $2 \frac{M_1^2}{\sigma^2} \|x_0 - \bar{x}\|^2$, para todo k . Além disso,

$$s_{k+1} - s_k = \frac{M_1^2}{\sigma^2} (\|x_0 - \bar{x}\|^2 - \|x_{k+2} - \bar{x}\|^2) - \frac{M_1^2}{\sigma^2} (\|x_0 - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2) = \frac{M_1^2}{\sigma^2} (\|x_{k+1} - \bar{x}\|^2 - \|x_{k+2} - \bar{x}\|^2).$$

Pela inequação (16) segue que $s_{k+1} - s_k \geq 0$ para todo k . Logo, a seqüência $\{s_k\}$ é monótona não-decrescente. Logo, a seqüência $\{s_k\}$ converge. Como $\sum_{k=0}^{\infty} \|x_k - z_k\|^4$ e $\sum_{k=0}^{\infty} \frac{M_1^2}{\sigma^2} \sum_{k=0}^{\infty} (\|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2)$ são séries de termos não-negativos. Pelo fato de $\sum_{k=0}^{\infty} \|x_k - z_k\|^4 < +\infty$, podemos usar o teste da comparação e afirmar que $\sum_{k=0}^{\infty} \|x_k - z_k\|^4 < +\infty$ e conseqüentemente $\|x_k - z_k\|^4 \rightarrow 0$ desta forma,

$$\lim_{k \rightarrow \infty} t_k \|d_k\| = \lim_{k \rightarrow \infty} \|x_k - z_k\| = 0,$$

de acordo com o fato acima, podemos concluir que,

$$\lim_{k \rightarrow \infty} (x_k - z_k) = 0.$$

□

Demonstração. (3) Note que,

$$\|x_k - x_{k+1}\| = \|x_k - P_\Omega[x_k - \alpha_k F(z_k)]\|,$$

como $x_0 \in \Omega$ e $x_k = P_\Omega[x_{k-1} - \alpha_{k-1} F(z_{k-1})]$ caso $k \geq 1$, podemos afirmar que $x_k \in \Omega$, conseqüentemente $P_\Omega[x_k] = x_k$, assim podemos desenvolver a igualdade acima usando a não-expansividade do operador projeção da seguinte forma,

$$\begin{aligned} \|x_k - x_{k+1}\| &= \|x_k - P_\Omega[x_k - \alpha_k F(z_k)]\| \\ &= \|P_\Omega[x_k] - P_\Omega[x_k - \alpha_k F(z_k)]\| \\ &\leq \|x_k - (x_k - \alpha_k F(z_k))\| \\ &= \|\alpha_k F(z_k)\| \\ &= \left\| \frac{F(z_k)^\top (x_k - z_k)}{\|F(z_k)\|^2} F(z_k) \right\| \\ &= \frac{|F(z_k)^\top (x_k - z_k)|}{\|F(z_k)\|} \\ &\leq \frac{\|F(z_k)\| \|x_k - z_k\|}{\|F(z_k)\|} \quad (\text{pela desigualdade de Cauchy-Schwarz}) \\ &= \|x_k - z_k\|. \end{aligned}$$

Ou seja,

$$0 \leq \|x_k - x_{k+1}\| \leq \|x_k - z_k\|.$$

Pelo item (2) desse Lema temos que $(x_k - z_k) \rightarrow 0$ assim, pela continuidade da norma segue que $\|x_k - z_k\| \rightarrow 0$. Pelo teorema do confronto obtemos que,

$$\lim_{k \rightarrow \infty} \|x_k - x_{k+1}\| = 0.$$

Portanto, $\lim_{k \rightarrow \infty} (x_k - x_{k+1}) = 0$. □

Agora, tem-se os resultados necessários para estabelecer a convergência global do Algoritmo 1.

Teorema 4.1 (Convergência Global). *Suponha que F é lipschitz contínua em Ω , continuamente diferenciável e monótona, além disso $\Phi \neq \emptyset$ e considerando as sequências $\{x_k\}$ e $\{d_k\}$ geradas pelo Algoritmo 1. Em tais condições, podemos afirmar que*

$$\liminf_{k \rightarrow \infty} \|F_k\| = 0. \quad (18)$$

Demonstração. Suponha por absurdo que (18) não seja válida. Então, existe um $\epsilon > 0$ tal que

$$\|F_k\| \geq \epsilon, \quad \forall k \geq 0. \quad (19)$$

Pelo Lema 3.1, usando o fato de um número real ser menor do que ou igual ao seu módulo e desigualdade de Cauchy-Schwarz, temos

$$\|F_k\| \|d_k\| \geq -F_k^\top d_k \geq c \|F_k\|^2,$$

o que implica, $\|F_k\| \|d_k\| \geq c \|F_k\|^2$ caso dividirmos ambos os membros desta desigualdade por $\|F_k\|$ e usarmos (19), obtemos

$$\|d_k\| \geq c \|F_k\| \geq c\epsilon, \quad \forall k \geq 0. \quad (20)$$

Pelo fato de $y_k = \gamma_k + \lambda_k t_k \|F_k\| d_k$ de modo que $\gamma_k = F_{k+1} - F_k$, $s_k = t_k d_k$, $t_k > 0$ e

$$\lambda_k = 1 + \|F_k\|^{-1} \max \left\{ 0, -\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} \right\}.$$

Temos

$$\langle y_k, t_k d_k \rangle = \langle \gamma_k + \lambda_k t_k \|F_k\| d_k, t_k d_k \rangle = \langle \gamma_k, t_k d_k \rangle + \langle \lambda_k t_k \|F_k\| d_k, t_k d_k \rangle = \langle \gamma_k, t_k d_k \rangle + \lambda_k \|F_k\| t_k^2 \langle d_k, d_k \rangle = \langle \gamma_k, t_k d_k \rangle + \lambda_k \|F_k\| \|t_k d_k\|^2.$$

Observe que se $-\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} \geq 0$ então,

$$\langle \gamma_k, t_k d_k \rangle + \lambda_k \|F_k\| \|t_k d_k\|^2 = \langle \gamma_k, t_k d_k \rangle + \left(1 - \|F_k\|^{-1} \frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} \right) \|F_k\| \|t_k d_k\|^2.$$

Caso, $-\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} < 0$, notamos que

$$\langle \gamma_k, t_k d_k \rangle + \lambda_k \|F_k\| \|t_k d_k\|^2 = \langle \gamma_k, t_k d_k \rangle > \langle \gamma_k, t_k d_k \rangle + \left(1 - \|F_k\|^{-1} \frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} \right) \|F_k\| \|t_k d_k\|^2.$$

Assim, podemos afirmar que independentemente do valor de $-\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2}$ é

$$\begin{aligned} \langle y_k, t_k d_k \rangle &= \langle \gamma_k, t_k d_k \rangle + \lambda_k \|F_k\| \|t_k d_k\|^2 \\ &\geq \langle \gamma_k, t_k d_k \rangle + \left(1 - \|F_k\|^{-1} \frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} \right) \|F_k\| \|t_k d_k\|^2 \\ &= \langle \gamma_k, t_k d_k \rangle + \|F_k\| \|t_k d_k\|^2 - \langle \gamma_k, t_k d_k \rangle \\ &= \|F_k\| \|t_k d_k\|^2 \\ &\geq \epsilon t_k^2 \|d_k\|^2. \end{aligned}$$

Como a desigualdade é válida para todo $k \geq 0$, $\langle y_{k-1}, t_{k-1}d_{k-1} \rangle \geq \epsilon t_{k-1}^2 \|d_{k-1}\|^2$ é uma desigualdade verdadeira, além disso pela igualdade

$$\langle y_{k-1}, t_{k-1}d_{k-1} \rangle = \langle t_{k-1}d_{k-1}, y_{k-1} \rangle = t_{k-1}d_{k-1}^\top y_{k-1}.$$

Obtemos a seguintes relação abaixo,

$$d_{k-1}^\top y_{k-1} = \epsilon t_{k-1} \|d_{k-1}\|^2, \quad (21)$$

ou equivalentemente,

$$s_{k-1}^\top y_{k-1} = \epsilon \|s_{k-1}\|^2. \quad (22)$$

Ademais, pelo fato de F ser Lipschitz contínua em Ω e aos resultados mostrados nos lemas anteriores temos

$$\begin{aligned} \|y_k\| &= \|\gamma_k + \lambda_k t_k \|F_k\| d_k\| \\ &\leq \|\gamma_k\| + \lambda_k \|F_k\| \|t_k d_k\| \\ &= \|\gamma_k\| + \left(1 + \|F_k\|^{-1} \max\left\{0, -\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2}\right\}\right) \|F_k\| \|t_k d_k\| \\ &\leq \|\gamma_k\| + \left(1 + \|F_k\|^{-1} \max\left\{0, \left|\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2}\right|\right\}\right) \|F_k\| \|t_k d_k\| \\ &= \|\gamma_k\| + \left(1 + \|F_k\|^{-1} \frac{|\langle \gamma_k, t_k d_k \rangle|}{\|t_k d_k\|^2}\right) \|F_k\| \|t_k d_k\| \\ &= \|\gamma_k\| + \|F_k\| \|t_k d_k\| + \frac{|\langle \gamma_k, t_k d_k \rangle|}{\|t_k d_k\|} \\ &\leq \|\gamma_k\| + \|F_k\| \|t_k d_k\| + \frac{\|\gamma_k\| \|t_k d_k\|}{\|t_k d_k\|} \\ &= 2\|\gamma_k\| + \|F_k\| \|t_k d_k\| \\ &= 2\|F_{k+1} - F_k\| + \|F_k\| \|t_k d_k\| \\ &\leq 2L\|x_{k+1} - x_k\| + \|F_k\| \|t_k d_k\| \\ &\leq 2L\|x_{k+1} - x_k\| + Mt_k \|d_k\| \\ &\leq 2L\|z_k - x_k\| + Mt_k \|d_k\| \\ &= 2Lt_k \|d_k\| + Mt_k \|d_k\| \\ &= (2L + M)t_k \|d_k\|. \end{aligned}$$

Ou seja,

$$\|y_{k-1}\| \leq (2L + M)t_{k-1} \|d_{k-1}\|. \quad (23)$$

O restante da prova é dividido em dois casos. Se $\beta_k(\tau_{k-1}) \geq \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2}$, então $\beta_k^+(\tau_{k-1}) = \beta_k(\tau_{k-1})$. Dessa forma,

$$\begin{aligned} |\beta_k(\tau_{k-1})| &= \left| \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left(\tau_{k-1} + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right| \\ &= \left| \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left(\left[\theta \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} + (1 - \theta) \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right] + \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right| \\ &= \left| \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - \left((\theta + 1) \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} - \theta \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \right) \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right| \\ &= \left| \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - (\theta + 1) \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} + \theta \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right| \\ &\leq \left| \frac{F_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} \right| + (\theta + 1) \left| \frac{\|y_{k-1}\|^2}{s_{k-1}^\top y_{k-1}} \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right| + \theta \left| \frac{s_{k-1}^\top y_{k-1}}{\|s_{k-1}\|^2} \frac{F_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}} \right| \\ &\leq \frac{|F_k^\top y_{k-1}|}{|d_{k-1}^\top y_{k-1}|} + 2 \frac{\|y_{k-1}\|^2}{|s_{k-1}^\top y_{k-1}|} \frac{|F_k^\top s_{k-1}|}{|d_{k-1}^\top y_{k-1}|} + 2 \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \frac{|F_k^\top s_{k-1}|}{|d_{k-1}^\top y_{k-1}|} \\ &\leq \frac{|F_k^\top y_{k-1}|}{|d_{k-1}^\top y_{k-1}|} + 2 \left(\frac{\|y_{k-1}\|^2}{|s_{k-1}^\top y_{k-1}|} + \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \right) \frac{|F_k^\top s_{k-1}|}{|d_{k-1}^\top y_{k-1}|} \\ &\leq \frac{\|F_k\| \|y_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} + 2 \left(\frac{\|y_{k-1}\|^2}{|s_{k-1}^\top y_{k-1}|} + \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \right) \frac{\|F_k\| \|s_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \\ &\leq \frac{M \|y_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} + 2 \left(\frac{\|y_{k-1}\|^2}{|s_{k-1}^\top y_{k-1}|} + \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \right) \frac{M \|s_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \\ &\leq \frac{M(2L + M)t_{k-1} \|d_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} + 2 \left(\frac{((2L + M)t_{k-1} \|d_{k-1}\|)^2}{|s_{k-1}^\top y_{k-1}|} + \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \right) \frac{M \|s_{k-1}\|}{|d_{k-1}^\top y_{k-1}|} \\ &\leq \frac{M(2L + M)t_{k-1} \|d_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} + 2 \left(\frac{((2L + M)t_{k-1} \|d_{k-1}\|)^2}{|s_{k-1}^\top y_{k-1}|} + \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \right) \frac{M \|s_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} \\ &\leq \frac{(2L + M)Mt_{k-1} \|d_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} + 2 \left(\frac{(2L + M)^2 t_{k-1}^2 \|d_{k-1}\|^2}{\epsilon \|s_{k-1}\|^2} + \frac{|s_{k-1}^\top y_{k-1}|}{\|s_{k-1}\|^2} \right) \frac{M \|s_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} \\ &\leq \frac{(2L + M)Mt_{k-1} \|d_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} + 2 \left(\frac{(2L + M)^2 t_{k-1}^2 \|d_{k-1}\|^2}{\epsilon \|s_{k-1}\|^2} + \frac{\|s_{k-1}\| \|y_{k-1}\|}{\|s_{k-1}\|^2} \right) \frac{M \|s_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} \\ &\leq \frac{(2L + M)Mt_{k-1} \|d_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} + 2 \left[\frac{(2L + M)^2 t_{k-1}^2 \|d_{k-1}\|^2}{\epsilon \|s_{k-1}\|^2} + \frac{(2L + M)t_{k-1} \|d_{k-1}\| \|s_{k-1}\|}{\|s_{k-1}\|^2} \right] \frac{M \|s_{k-1}\|}{\epsilon t_{k-1} \|d_{k-1}\|^2} \\ &= \frac{(2L + M)M}{\epsilon \|d_{k-1}\|} + \frac{2(2L + M)^2 M}{\epsilon^2 \|d_{k-1}\|} + \frac{2(2L + M)M}{\epsilon \|d_{k-1}\|} \\ &= \left[\frac{3(2L + M)M}{\epsilon} + \frac{2(2L + M)^2 M}{\epsilon^2} \right] \frac{1}{\|d_{k-1}\|}. \end{aligned}$$

Denotando

$$\bar{M} = \frac{3(2L+M)M}{\varepsilon} + \frac{2(2L+M)^2M}{\varepsilon^2}.$$

Podemos escrever que

$$|\beta_k(\tau_{k-1})| \leq \frac{\bar{M}}{\|d_{k-1}\|}. \quad (24)$$

Caso $\beta_k(\tau_{k-1}) < \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2}$, segue que $\beta_k^+(\tau_{k-1}) = \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2}$.

Para $\eta \in [0, 1)$ temos,

$$|\beta_k^+(\tau_{k-1})| = \left| \eta \frac{F_k^\top d_{k-1}}{\|d_{k-1}\|^2} \right| \leq \eta \frac{\|F_k\| \|d_{k-1}\|}{\|d_{k-1}\|^2} \leq \frac{\eta M}{\|d_{k-1}\|} \leq \frac{M}{\|d_{k-1}\|},$$

o que significa

$$|\beta_k^+(\tau_{k-1})| \leq \frac{M}{\|d_{k-1}\|}. \quad (25)$$

Assim, podemos afirmar

$$\|d_k\| = \|-F_k + \beta_k^+(\tau_{k-1})d_{k-1}\| \leq \|-F_k\| + |\beta_k^+(\tau_{k-1})| \|d_{k-1}\| = \|F_k\| + |\beta_k^+(\tau_{k-1})| \|d_{k-1}\| \leq M + |\beta_k^+(\tau_{k-1})| \|d_{k-1}\|.$$

Note que $|\beta_k(\tau_{k-1})| \leq \frac{\bar{M}}{\|d_{k-1}\|}$ com $\bar{M} = \frac{3(2L+M)M}{\varepsilon} + \frac{2(2L+M)^2M}{\varepsilon^2}$ ou $|\beta_k^+(\tau_{k-1})| \leq \frac{M}{\|d_{k-1}\|}$. Logo,

$$\|d_k\| \leq M + |\beta_k^+(\tau_{k-1})| \|d_{k-1}\| \leq M + \frac{\bar{M}}{\|d_{k-1}\|} \|d_{k-1}\| = M + \bar{M}.$$

Ou,

$$\|d_k\| \leq M + |\beta_k^+(\tau_{k-1})| \|d_{k-1}\| \leq M + \frac{M}{\|d_{k-1}\|} \|d_{k-1}\| = M + M.$$

Com isso, obtemos

$$\|d_k\| \leq M + |\beta_k^+(\tau_{k-1})| \|d_{k-1}\| \leq \tilde{M}, \quad (26)$$

com $\tilde{M} = M + \max\{M, \bar{M}\}$.

No Lema 4.1 vimos que

$$t_k \geq \min \left\{ \xi, \frac{c\rho}{(L+\sigma)} \frac{\|F_k\|^2}{\|d_k\|^2} \right\}.$$

Usando as desigualdades $\|F_k\| \geq \varepsilon$, $\|d_k\| \geq c\varepsilon$ e $\tilde{M} \geq \|d_k\|$ verificamos a seguinte desigualdade

$$\begin{aligned} t_k \|d_k\| &\geq \min \left\{ \xi, \frac{c\rho}{(L+\sigma)} \frac{\|F_k\|^2}{\|d_k\|^2} \right\} \|d_k\| \\ &\geq \min \left\{ \xi, \frac{c\rho}{(L+\sigma)} \frac{\varepsilon^2}{\|d_k\|^2} \right\} \|d_k\| \\ &= \min \left\{ \xi \|d_k\|, \frac{c\rho}{(L+\sigma)} \frac{\varepsilon^2}{\|d_k\|} \right\} \\ &\geq \min \left\{ \xi \|d_k\|, \frac{c\rho\varepsilon^2}{(L+\sigma)\tilde{M}} \right\} \\ &\geq \min \left\{ \xi c\varepsilon, \frac{c\rho\varepsilon^2}{(L+\sigma)\tilde{M}} \right\}. \end{aligned}$$

Observe que com isso, se a sequência $t_k \|d_k\|$ convergisse teríamos que $t_k \|d_k\| \rightarrow \alpha \geq \min \left\{ \xi c\varepsilon, \frac{c\rho\varepsilon^2}{(L+\sigma)\tilde{M}} \right\} > 0$. Isso contradiz o item (2) do Lema 4.4. Consequentemente não existe um $\varepsilon > 0$ tal que $\|F_k\| \geq \varepsilon$ e portanto $\liminf_{k \rightarrow \infty} \|F_k\| = 0$. \square

5 Taxa de Convergência

Vimos que a sequência $\{x_k\}$ gerada pelo Algoritmo 1 converge globalmente para a solução do problema. Agora só nos resta saber em que taxa essa convergência ocorre.

Os autores mostraram que tal sequência converge linearmente desde que seja assumido a seguinte hipótese: para cada $x \in \Phi$, existem as constantes $\mu \in (0, 1)$ e $\nu > 0$ tais que

$$\mu \operatorname{dist}(x, \Phi) \leq \|F(x)\|, \quad \forall x \in \mathcal{N}_\nu(\bar{x}), \quad (27)$$

onde $\mathcal{N}_\nu(\bar{x})$ é a vizinhança de \bar{x} definida por $\mathcal{N}_\nu(\bar{x}) = \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \nu\}$. Com isso, podemos propor o resultado subsequente.

Teorema 5.1. *Suponha que F é lipschitz contínua em Ω , continuamente diferenciável e monótona além disso, $\Phi \neq \emptyset$ e por fim, existem as constantes $\mu \in (0, 1)$ e $\nu > 0$ tais que (27) é válida. Se a sequência $\{x_k\}$ é gerada pelo Algoritmo 1, então a sequência gerada converge linearmente.*

Demonstração. Seja $\mathcal{Q}_k = \arg \min\{\|x_k - x\| : x \in \Phi\}$, isto é, $\|x_k - \mathcal{Q}_k\| = \operatorname{dist}(x_k, \Phi)$. Relembrando, no Lema 4.2, deduzimos a inequação (15), dada por

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \frac{[F(z_k)^\top (x_k - z_k)]^2}{\|F(z_k)\|^2}.$$

Consequentemente,

$$\begin{aligned}
\text{dist}(x_{k+1}, \Phi)^2 &= \|x_{k+1} - \mathcal{Q}_{k+1}\|^2 \\
&\leq \|x_{k+1} - \mathcal{Q}_k\|^2 \\
&\leq \|x_k - \mathcal{Q}_k\|^2 - \frac{[F(z_k)^\top(x_k - z_k)]^2}{\|F(z_k)\|^2} \quad (\text{por (15)}) \\
&= \text{dist}(x_k, \Phi)^2 - \frac{[F(z_k)^\top(x_k - z_k)]^2}{\|F(z_k)\|^2} \\
&= \text{dist}(x_k, \Phi)^2 - \frac{[-F(z_k)^\top t_k d_k]^2}{\|F(z_k)\|^2} \\
&= \text{dist}(x_k, \Phi)^2 - \frac{t_k^2 [-F(z_k)^\top d_k]^2}{\|F(z_k)\|^2}.
\end{aligned}$$

Pelo Lema 3.1 existe uma constante $c \in (0, 1)$ tal que $F(z_k)^\top d_k \leq -c\|F(z_k)\|^2$ logo,

$$\text{dist}(x_{k+1}, \Phi)^2 \leq \text{dist}(x_k, \Phi)^2 - \frac{t_k^2 [-F(z_k)^\top d_k]^2}{\|F(z_k)\|^2} \leq \text{dist}(x_k, \Phi)^2 - \frac{t_k^2 [c\|F(z_k)\|^2]^2}{\|F(z_k)\|^2} = \text{dist}(x_k, \Phi)^2 - c^2 t_k^2 \|F(z_k)\|^2.$$

Usando a busca linear (10) podemos obter as seguintes desigualdades,

$$\|F(z_k)\| \|d_k\| \geq |F(z_k)^\top d_k| \geq -F(z_k)^\top d_k \geq \sigma t_k \|d_k\|^2.$$

Logo, $\|F(z_k)\| \geq \sigma t_k \|d_k\|$. E pelo Lema 3.1

$$\|F_k\| \|d_k\| \geq |F_k^\top d_k| \geq -F_k^\top d_k \geq c \|F_k\|^2.$$

Daí, $\|d_k\| \geq c \|F_k\|$. Desta forma,

$$\begin{aligned}
\text{dist}(x_{k+1}, \Phi)^2 &\leq \text{dist}(x_k, \Phi)^2 - c^2 t_k^2 \|F(z_k)\|^2 \\
&\leq \text{dist}(x_k, \Phi)^2 - \sigma^2 c^2 t_k^4 \|d_k\|^2 \\
&\leq \text{dist}(x_k, \Phi)^2 - \sigma^2 c^4 t_k^4 \|F_k\|^2 \\
&\leq \text{dist}(x_k, \Phi)^2 - \sigma^2 \mu^2 c^4 t_k^4 \text{dist}(x_k, \Phi)^2 \\
&= (1 - \sigma^2 \mu^2 c^4 t_k^4) \text{dist}(x_k, \Phi)^2.
\end{aligned}$$

Note que $\sigma, \mu, c, t_k \in (0, 1)$. Portanto, a sequência $\{x_k\}$ converge linearmente. \square

6 Experimentos Numéricos

Experimentos numéricos foram realizados para estudar o comportamento do Algoritmo 1 mediante a variação dos valores dos parâmetros η e θ , presentes nas equações (5) e (8) responsáveis pela determinação da direção conjugada definida em (4). Para isso tomamos as seguintes funções disponíveis na biblioteca virtual de funções testes da Simon Fraser University disponível em <https://www.sfu.ca/~ssurjano/optimization.html>: SUM SQUARES FUNCTION e TRID FUNCTION.

Por conveniência, chamaremos SUM SQUARES FUNCTION de f e TRIDFUNCTION de g . Onde $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ tais que $f(x) = \sum_{i=1}^n ix_i^2$ e $g(x) = \sum_{i=1}^n (x_i - 1)^2 - \sum_{i=2}^n x_i x_{i-1}$. Evidentemente, tais funções não são campos vetoriais mas, como aprendemos em Cálculo, os pontos críticos de uma função são os elementos do domínio que anulam o campo gradiente e onde o gradiente não é bem definido. Sabemos que os pontos críticos são candidatos a mínimo local da função (se existir) além disso, o site nos fornece que f, g são funções com um único mínimo local em todo o domínio, ou seja, possuem um único ponto crítico do qual possivelmente é aquele que anula o campo gradiente.

Dito isso, temos

$$\frac{\partial f}{\partial x_i}(x) = \frac{\partial}{\partial x_i} \left(\sum_{i=1}^n ix_i^2 \right) = 2ix_i \text{ para todo } 1 \leq i \leq n \text{ com } i \in \mathbb{N}.$$

Além disso, note que

$$\frac{\partial g}{\partial x_1}(x) = \frac{\partial}{\partial x_1} \left(\sum_{i=1}^n (x_i - 1)^2 - \sum_{i=2}^n x_i x_{i-1} \right) = \frac{\partial}{\partial x_1} ((x_1 - 1)^2 - x_2 x_1) = 2(x_1 - 1) - x_2 = 2x_1 - x_2 - 2,$$

para $1 < i < n$ segue se

$$\frac{\partial g}{\partial x_i}(x) = \frac{\partial}{\partial x_i} \left(\sum_{i=1}^n (x_i - 1)^2 - \sum_{i=2}^n x_i x_{i-1} \right) = \frac{\partial}{\partial x_i} ((x_i - 1)^2 - x_i x_{i-1} - x_{i+1} x_i) = 2(x_i - 1) - x_{i-1} - x_{i+1} = 2x_i - x_{i-1} - x_{i+1} - 2,$$

por fim,

$$\frac{\partial g}{\partial x_n}(x) = \frac{\partial}{\partial x_n} \left(\sum_{i=1}^n (x_i - 1)^2 - \sum_{i=2}^n x_i x_{i-1} \right) = \frac{\partial}{\partial x_n} ((x_n - 1)^2 - x_n x_{n-1}) = 2(x_n - 1) - x_{n-1} = 2x_n - x_{n-1} - 2.$$

Assim, temos os campos gradientes $\nabla f, \nabla g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tais que

$$\nabla f(x_1, x_2, \dots, x_i, \dots, x_n) = (2x_1, 4x_2, \dots, 2ix_i, \dots, 2nx_n),$$

$$\nabla g(x_1, x_2, \dots, x_i, \dots, x_n) = (2x_1 - x_2 - 2, 2x_2 - x_3 - 2, \dots, 2x_i - x_{i-1} - x_{i+1} - 2, \dots, 2x_n - x_{n-1} - 2).$$

Novamente, por conveniência, chamaremos ∇f de F e ∇g de G . Observe que caso F e G possuam zeros, encontrá-los é suficiente para minimizar as funções f e g . Por isso, verificaremos se os campos gradientes, satisfazem as hipóteses do

Teorema 4.1. Primeiramente, note que $F(0) = (2 \cdot 0, 4 \cdot 0, \dots, 2i \cdot 0, \dots, 2n \cdot 0) = 0$, ou seja $\Phi \neq \emptyset$. Além disso, considerando $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$ temos que

$$\begin{aligned} F(x) - F(y) &= (2x_1, 4x_2, \dots, 2ix_i, \dots, 2nx_n) - (2y_1, 4y_2, \dots, 2iy_i, \dots, 2ny_n) \\ &= (2(x_1 - y_1), 4(x_2 - y_2), \dots, 2i(x_i - y_i), \dots, 2n(x_n - y_n)). \end{aligned}$$

Consequentemente,

$$\begin{aligned} (F(x) - F(y))^T(x - y) &= (2(x_1 - y_1), 4(x_2 - y_2), \dots, 2i(x_i - y_i), \dots, 2n(x_n - y_n))^T(x_1 - y_1, x_2 - y_2, \dots, x_i - y_i, \dots, x_n - y_n) \\ &= 2(x_1 - y_1)^2 + 4(x_2 - y_2)^2 + \dots + 2i(x_i - y_i)^2 + \dots + 2n(x_n - y_n)^2 \\ &= \sum_{i=1}^n 2i(x_i - y_i)^2 \\ &\geq 0. \end{aligned}$$

Como x e y eram vetores arbitrários de \mathbb{R}^n conclui-se que a desigualdade acima é válida para quaisquer $x, y \in \mathbb{R}^n$. Por fim,

$$\begin{aligned} \|F(x) - F(y)\| &= \sqrt{(F(x) - F(y))^T(F(x) - F(y))} \\ &= \sqrt{4(x_1 - y_1)^2 + 16(x_2 - y_2)^2 + \dots + 4i^2(x_i - y_i)^2 + \dots + 4n^2(x_n - y_n)^2} \\ &= \sqrt{\sum_{i=1}^n 4i^2(x_i - y_i)^2} \\ &\leq \sqrt{\sum_{i=1}^n 4n^2(x_i - y_i)^2} \\ &= 2n \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &= 2n\|x - y\|. \end{aligned}$$

Note que $2n > 0$ assim, tomando $L = 2n$, temos a existência de uma constante $L > 0$ tal que $\|F(x) - F(y)\| \leq L\|x - y\|$. Como x e y eram vetores arbitrários de \mathbb{R}^n , concluímos que essa desigualdade é válida para quaisquer $x, y \in \mathbb{R}^n \supset \Omega$. Portanto, F satisfaz as três hipóteses do Teorema 4.1.

Referências

- [1] Dai YH, Kou CX. A nonlinear conjugate gradient algorithm with an optimal property and an improved wolfe line search. *SIAM J Optim.*, 2013;23:296-320.
- [2] Yanyun Ding, Yunhai Xiao e Jianwei Li. A class of conjugate gradient methods for convex constrained monotone equations. Em: **Optimization** 66.12 (2017), pp. 2309-2328.
- [3] Alexey Izmailov e Mikhhail Solodov. **Otimização - volume 2: métodos computacionais**. Terceira Edição. IMPA, 2018. ISBN: 978-85-244-0454-2.