

LISTA DE EXERCÍCIOS 2
MAC0427 PROGRAMAÇÃO NÃO LINEAR
ALGORITMOS PARA OTIMIZAÇÃO IRRESTRITA

Entrega: somente dos exercícios 3, 5, 10 e 13, nos primeiros 15 minutos da aula de 05/05.

Exercícios adaptados dos livros listados como “Material para estudo” no PACA.

Exercício 1. Seja $f: \mathbb{R}^n \rightarrow \mathbb{R}$ continuamente diferenciável e seja $\bar{x} \in \mathbb{R}^n$ tal que $\nabla f(\bar{x}) \neq 0$. Seja $M \in \mathbb{R}^{n \times n}$ uma matriz simétrica positiva definida. Mostre que $d := -M\nabla f(\bar{x})$ é uma direção de descida para f em \bar{x} .

Exercício 2. Seja $(\alpha_t)_{t=0}^\infty$ uma sequência de reais positivos. Suponha que o limite $L := \lim_{t \rightarrow \infty} \frac{\alpha_{t+1}}{\alpha_t}$ existe. Mostre que, se $L \in (0, 1)$, então $(\alpha_t)_{t=0}^\infty$ converge linearmente para 0. Mostre também que, se $L = 0$, então $(\alpha_t)_{t=0}^\infty$ converge superlinearmente para 0.

Exercício 3. Mostre que a sequência $(\alpha_t)_{t=1}^\infty$ dada por $\alpha_t := 1/t$ converge sublinearmente para 0.

Exercício 4. Defina $\varepsilon_t := 500(1 - \frac{t}{3t+1})^t$ para todo inteiro $t \geq 0$. A sequência $(\varepsilon_t)_{t=0}^\infty$ converge linearmente para 0. Prove que a razão de convergência é $2/3$.

Exercício 5. Considere a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Encontre expressões para $\phi(\alpha) := f(0 + \alpha d)$ para as direções de descida $d \in \mathbb{R}^2$ dos seguintes métodos:

- (i) método do gradiente;
- (ii) método de Newton;
- (iii) método de quase Newton, substituindo o hessiano de f em 0 pela matriz

$$\begin{bmatrix} 1/2 & 0 \\ 0 & 200 \end{bmatrix}.$$

Mostre seus passos intermediários.

Exercício 6. Se $f(x) = \frac{1}{2}x^T Ax + b^T x$, com $A \succ 0$, e $\phi(\alpha) := f(\bar{x} + \alpha d)$ para certos $\bar{x}, d \in \mathbb{R}^n$ com $d \neq 0$, mostre que o valor de α que minimiza $\phi(\alpha)$ sobre $\alpha \in \mathbb{R}$ é

$$\alpha^* = -\frac{\nabla f(\bar{x})^T d}{d^T \nabla^2 f(\bar{x}) d}.$$

Exercício 7. Considere a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $f(x) = (x_1 + x_2^2)^2$. Vamos fazer uma busca direcional em f a partir do ponto $\bar{x} := (1, 0)^T$ na direção $d := (-1, 1)^T$, isto é,

$$\phi(\alpha) := f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right)$$

para $\alpha \geq 0$. Mostre que d é uma direção de descida a partir de \bar{x} e encontre todas as soluções ótimas para a busca linear exata correspondente, isto é, encontre todas as solução ótimas para o problema de minimizar $\phi(\alpha)$ sujeito a $\alpha \geq 0$.

Exercício 8. Considere a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x) = (x_1 - x_2)^2 + \frac{(1 - x_1)^2}{2}.$$

Numa iteração do método de busca direcional a partir do ponto $\bar{x} := (0, -1)^\top$ e na direção $d := (\frac{1}{2}, 1)^\top$, encontre um passo $\alpha > 0$ satisfazendo a condição de Armijo com o parâmetro $1/8$. Seu passo deve ser “de ordem maximal” com razão $\eta := 2$, isto é, α satisfaz Armijo mas $\eta\alpha$ não. Utilize o algoritmo visto em aula com passo inicial $\alpha_0 := 1$.

Exercício 9. Considere a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x) = x_1^4 - 3x_1^3 + 4x_1x_2^2,$$

numa iteração do método de busca direcional a partir do ponto $\bar{x} := (3, -1)^\top$ e na direção $d := (-1, 1)^\top$. Com relação às condições de Wolfe com os parâmetros $1/5$ para Armijo e $4/5$ para curvatura, determine:

- (i) se a direção é de descida;
- (ii) se o passo $\alpha = 1$ satisfaz as condições de Wolfe;
- (iii) se o passo $\alpha = 2$ satisfaz as condições de Wolfe;
- (iv) se o passo $\alpha = 5$ satisfaz as condições de Wolfe;
- (v) se o passo $\alpha = 7$ satisfaz as condições de Wolfe;
- (vi) se o passo $\alpha = 7$ satisfaz as condições fortes de Wolfe.

Exercício 10. Ao usar o método do gradiente com busca linear exata para minimizar a função quadrática $f(x) = \frac{1}{2}x^\top Ax - b^\top x$ com $A \succ 0$, mostre que duas direções de descida consecutivas d_t e d_{t+1} são ortogonais. (Você pode usar o fato, pedido em outro exercício, de que o tamanho do passo é dado por

$$\alpha_t = \frac{\nabla f_t^\top \nabla f_t}{\nabla f_t^\top A \nabla f_t},$$

onde $\nabla f_t := \nabla f(x_t)$.)

Exercício 11. Ao usar o método do gradiente com busca linear exata para minimizar uma função $f: \mathbb{R}^n \rightarrow \mathbb{R}$ continuamente diferenciável, mostre que duas direções de descida consecutivas d_t e d_{t+1} são ortogonais.

Exercício 12. Considere a função $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x) = \frac{1}{2}(x_1^2 - x_2)^2 + \frac{1}{2}(1 - x_1)^2.$$

Mostre que a direção de Newton a partir do ponto $\bar{x} := (0, 1)^\top$ não é uma direção de descida (basta mostrar que $\phi(\alpha) \geq \phi(0)$ para todo $\alpha > 0$ suficientemente pequeno, para a definição apropriada de ϕ). “Corrija” a hessiana de f em \bar{x} através de sua decomposição espectral de modo que todas entradas da matriz diagonal sejam $\geq 1/2$. Mostre que a direção “corrigida” é de descida (basta mostrar que $\phi'(0) < 0$ para uma outra definição apropriada de ϕ).

Exercício 13. Considere que vamos minimizar a função $f(x) := x^4$ utilizando o método de Newton a partir de um ponto arbitrário $x_0 \in \mathbb{R} \setminus \{0\}$. Calcule a fórmula de atualização de x_{t+1} em termos de x_t (com passo unitário), e determine a velocidade de convergência da sequência $(|x_t - x^*|)_{t=0}^\infty$, onde x^* é o único minimizador global de f .

Exercício 14. Considere que vamos minimizar a função $f(x) := \sqrt{1 + x^2}$ utilizando o método de Newton a partir de um ponto arbitrário $x_0 \in \mathbb{R} \setminus \{0\}$. Calcule a fórmula de atualização de x_{t+1} em termos de x_t (com passo unitário) e determine a velocidade de convergência da sequência $(|x_t - x^*|)_{t=0}^\infty$, onde x^* é o único minimizador global de f . Note que a velocidade pode variar de acordo com a escolha de x_0 .

Exercício 15. Considere utilizar o método do gradiente para encontrar um minimizador global de $f: x \in \mathbb{R}^2 \mapsto \frac{1}{2}x_1^2 + \frac{1}{4}x_2^4 - \frac{1}{2}x_2^2$ a partir de $x_0 := e_1$. Determine todos os minimizadores globais de f e mostre que o método do gradiente não converge para nenhum desses minimizadores.

Exercício 16. Para encontrar um minimizador local de

$$f(x) := -\frac{x^4}{4} + \frac{x^3}{3} + x^2,$$

você aplica o método de Newton a partir do ponto $x_0 := 1$. Você decide que o tamanho de passo $\alpha = 1$ é satisfatório, uma vez que $x_1 := x_0 + d$, onde d é a direção de Newton, satisfaz $f'(x_1) = 0$. Explique por que essa escolha de passo não é boa.

Exercício 17. Seja $A \in \mathbb{R}^{n \times n}$ uma matriz inversível e sejam $u, v \in \mathbb{R}^n$. Verifique que, se $v^\top A^{-1}u \neq -1$, então

$$(A + uv^\top)^{-1} = A^{-1} + \frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u}.$$

Mostre também que, se $v^\top A^{-1}u = -1$, então a matriz $A + uv^\top$ não é inversível.

Exercício 18. Prove que, se $A \succ 0$ e $u \in \mathbb{R}^n$, então

$$\det(A + uu^\top) = (1 + u^\top A^{-1}u) \det(A). \quad (1)$$

Roteiro:

- (1) Primeiro prove (1) no caso especial em que $A = I$ e $\|u\|^2 = 1$; para tanto, use o fato de que, se $\|u\|^2 = 1$, então existe uma matriz *ortogonal* Q tal que $u = Qe_1$.
- (2) Repita o argumento anterior para o caso especial em que $A = I$ e $\|u\| > 0$. A mudança é que, agora, $u/\|u\| = Qe_1$.
- (3) Para remover a restrição de que $A = I$ do caso anterior, coloque em evidência alguns fatores *comuns* da expressão $A + uu^\top$; a saber, um fator (*comum!*) de $A^{1/2}$ à esquerda e outro fator (*comum!*) de $A^{1/2}$ à direita. Isso deve reduzir o caso geral ao caso em que $A = I$.

Exercício 19. Considere o método de quase Newton com atualização BFGS. Suponha que a estimativa C_t da inversa do hessiano é positiva definida e que $(\nabla f_{t+1} - \nabla f_t)^\top (x_{t+1} - x_t) > 0$. Mostre que $C_{t+1} \succ 0$.

Lembrete:

$$C_{t+1} = (I - \gamma_t s_t y_t^\top) C_t (I - \gamma_t y_t s_t^\top) + \gamma_t s_t s_t^\top,$$

onde

$$\begin{aligned} s_t &:= x_{t+1} - x_t, \\ y_t &:= \nabla f_{t+1} - \nabla f_t, \\ \gamma_t &:= \frac{1}{y_t^\top s_t}. \end{aligned}$$