UNIVERSIDADE ESTADUAL DE MONTES CLAROS - UNIMONTES
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS – CCET
HELLEN MICHELE PARAISO NERI
MATRÍCULA: 100015057

Conforme o "problema A", represente as curvas de nível e as iterações ocorridas, simultaneamente em um mesmo plano 2D x1-x2, utilizando qualquer software de sua escolha, simulando se utilizado os métodos de Newton e Gradiente para a sua solução. (Anexe o desenvolvimento no classroom) *

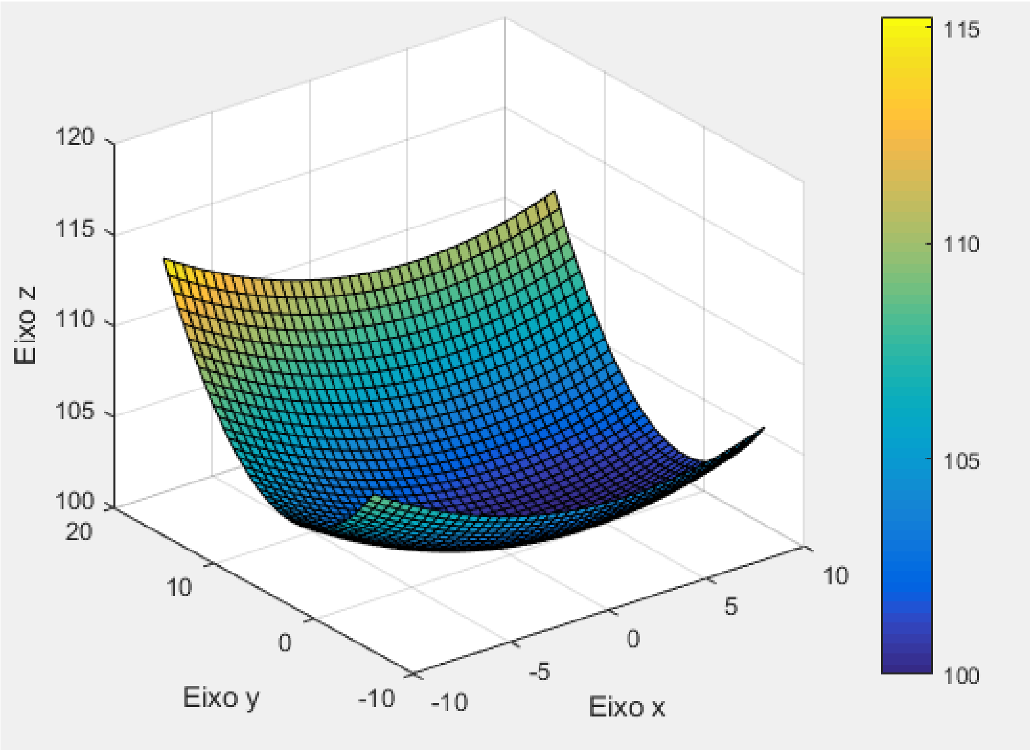
4 pontos

Considere X0 = (6,6) para o método do Gradiente, e X0=(10,15) para o método de Newton.

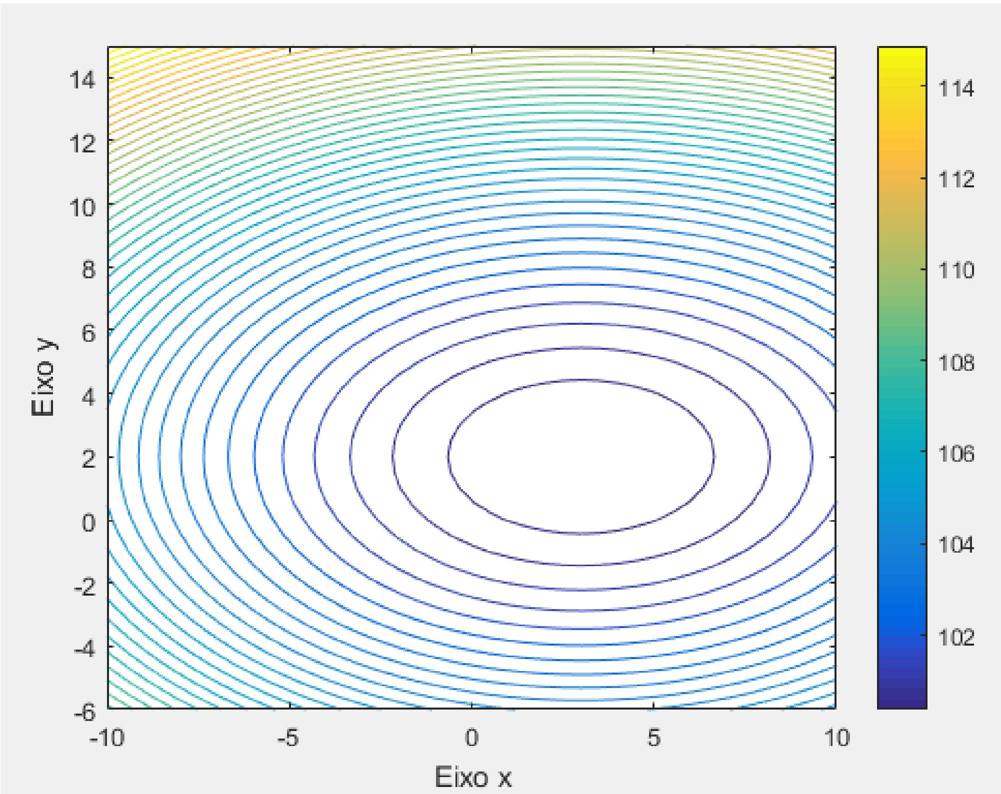
Código para geração das curvas de nível e do gráfico da função:

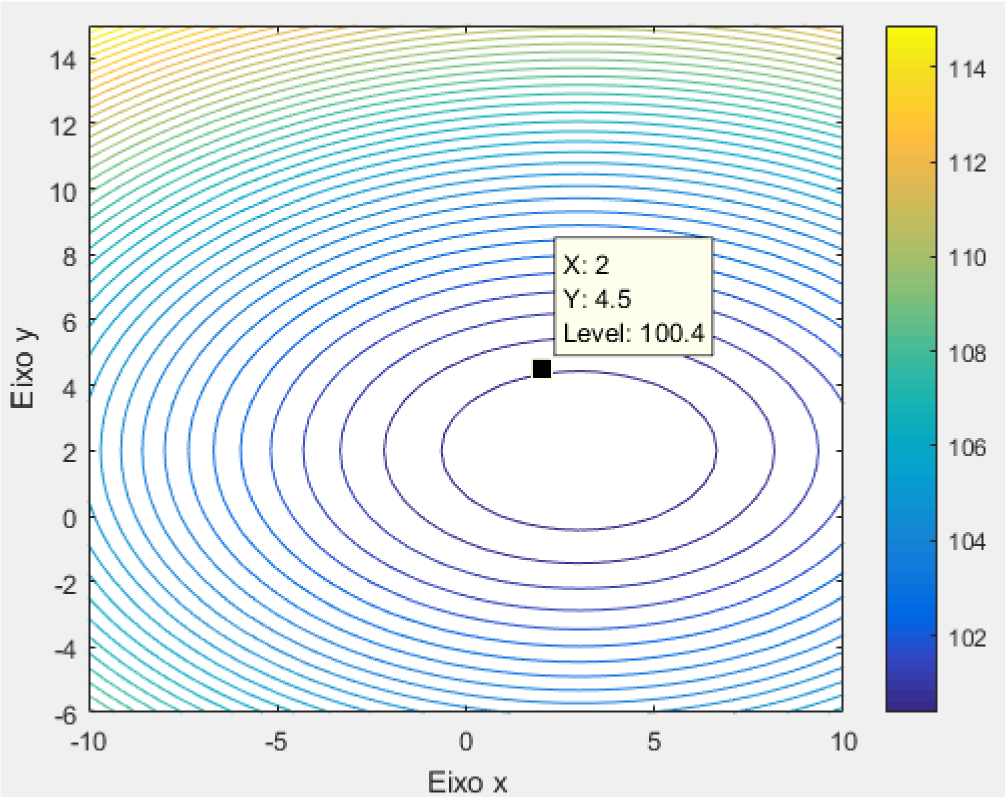
```
[X, Y] = meshgrid(-10:0.5:10, -6:0.5:15);
Z = (((X-3)./(6))).^(2)+(((Y-2)./(4))).^(2)+100;
figure
surf(X,Y,Z)
colorbar
xlabel('Eixo x')
ylabel('Eixo y')
zlabel('Eixo z')

figure
mesh(peaks)
pcolor(X,Y,Z)
contour(X,Y,Z,40)
colorbar
xlabel('Eixo x')
ylabel('Eixo y')
```

Gráfico da função

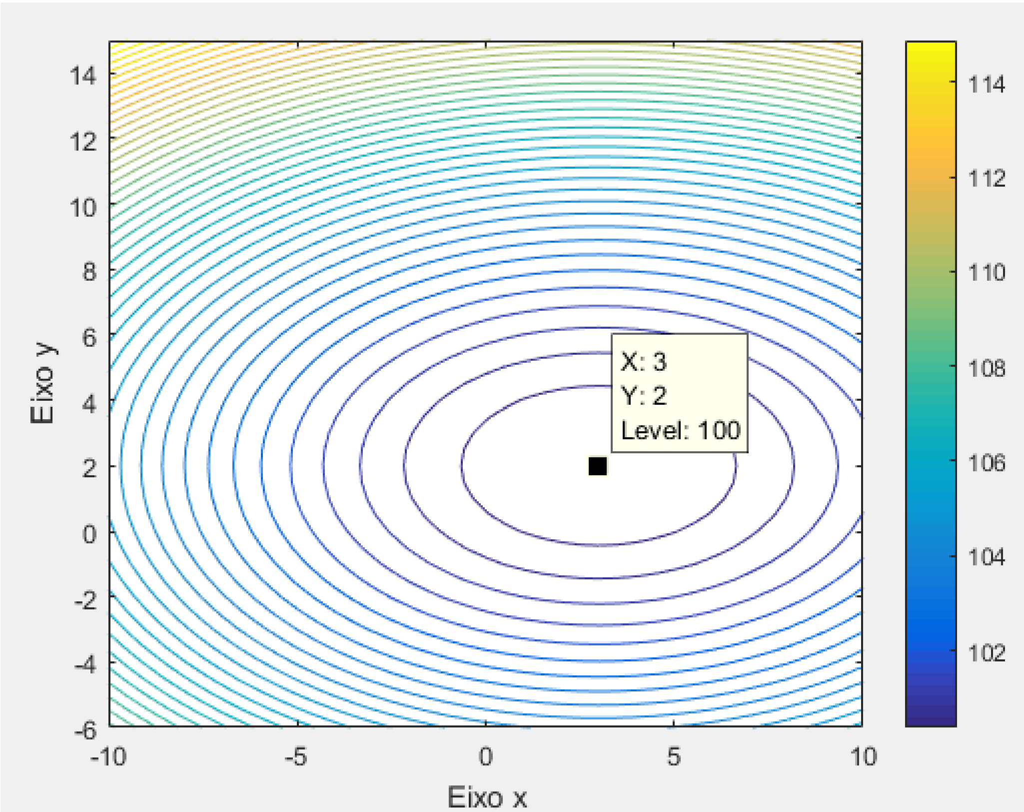


Curvas de nível da função





Ponto de mínimo global da função



Algoritmo do método do gradiente feito no MATLAB:

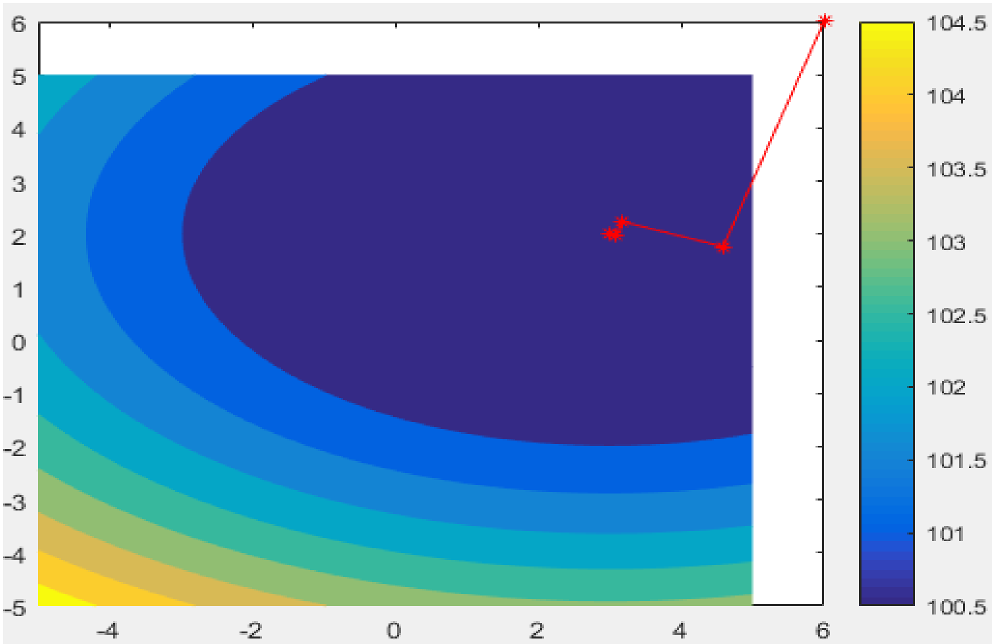
```
clc
clear
format long

% Function Definition (Enter your Function here):
syms X Y;
f = (((X-3)/(6)))^(2)+(((Y-2)/(4)))^(2)+100;

% Initial Guess:
x(1) = 6;
y(1) = 6;
e = 10^(-8); % Convergence Criteria
i = 1; % Iteration Counter

% Gradient Computation:
df_dx = diff(f, X);
df_dy = diff(f, Y);
J = [subs(df_dx,[X,Y], [x(1),y(1)]) subs(df_dy, [X,Y], [x(1),y(1)])];
% Gradient
S = -(J); % Search Direction

% Minimization Condition:
while norm(J) > e
    I = [x(i),y(i)]';
    syms h; % Step size
    g = subs(f, [X,Y], [x(i)+S(1)*h,y(i)+h*S(2)]);
    dg_dh = diff(g,h);
    h = solve(dg_dh, h); % Optimal Step Length
    x(i+1) = I(1)+h*S(1); % Updated x value
    y(i+1) = I(2)+h*S(2); % Updated y value
    i = i+1;
    J = [subs(df_dx,[X,Y], [x(i),y(i)]) subs(df_dy, [X,Y],
[x(i),y(i)])]; % Updated Gradient
    S = -(J); % New Search Direction
end

% Result Table:
Iter = 1:i;
X_coordinate = x';
Y_coordinate = y';
Iterations = Iter';
T = table(Iterations,X_coordinate,Y_coordinate);

% Plots:
fcontour(f, 'Fill', 'On');
hold on;
plot(x,y,'*-r');
colorbar

% Output:
fprintf('Valor da função objetivo inicial: %d\n\n',subs(f,[X,Y],
[x(1),y(1)]));
if (norm(J) < e)
    fprintf('Mínimo obtido com sucesso...\n\n');
end
fprintf('Número de Iterações para Convergência: %d\n\n', i);
fprintf('Ponto de Mínimo: : [%d,%d]\n\n', x(i), y(i));
```

```
fprintf('Valor mínimo da função objetivo após a otimização:: %d\n\n',
subs(f,[X,Y], [x(i),y(i)]));
disp(T);
```

Output:

```
Command Window
Valor da função objetivo inicial: 101

Mínimo obtido com sucesso...

Número de Iterações para Convergência: 14

Ponto de Mínimo: : [3.000000e+00,2.000000e+00]

Valor mínimo da função objetivo após a otimização:: 100

    Iterations      X_coordinate        Y_coordinate
    __________    ________________    ________________

        1                        6                   6
        2         4.58823529411765    1.76470588235294
        3         3.17647058823529    2.23529411764706
        4         3.09342560553633    1.98615916955017
        5         3.01038062283737    2.01384083044983
        6         3.00549562385508    1.99918583350295
        7         3.00061062487279    2.00081416649705
        8         3.00032327199148    1.99995210785311
        9         3.00003591911016    2.00004789214689
       10          3.0000190159995    1.99999718281489
       11         3.00000211288883    2.00000281718511
       12         3.00000111858821    1.99999983428323
       13         3.00000012428758    2.00000016571677
fx     14         3.00000006579931    1.99999999025195
```

Representação gráfico do método do gradiente



Algoritmo do método de Newton

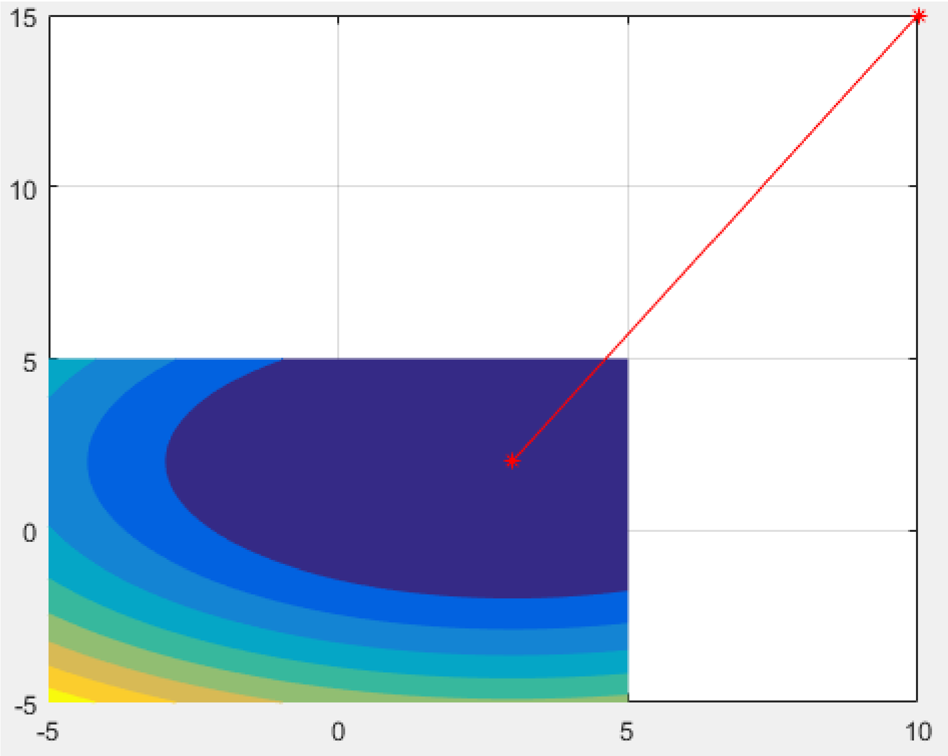
```
% Newton's Method

clc
clear
format long

% Function Definition (Enter your Function here):
syms X Y;
f =(((X-3)/(6)))^(2)+(((Y-2)/(4)))^(2)+100;

% Initial Guess (Choose Initial Guesses):
x(1) = 10;
y(1) = 15;
e = 10^(-8); % Convergence Criteria
i = 1; % Iteration Counter

% Gradient and Hessian Computation:
df_dx = diff(f, X);
df_dy = diff(f, Y);
J = [subs(df_dx,[X,Y], [x(1),y(1)]) subs(df_dy, [X,Y], [x(1),y(1)])];
% Gradient
ddf_ddx = diff(df_dx,X);
ddf_ddy = diff(df_dy,Y);
ddf_dxdy = diff(df_dx,Y);
ddf_ddx_1 = subs(ddf_ddx, [X,Y], [x(1),y(1)]);
ddf_ddy_1 = subs(ddf_ddy, [X,Y], [x(1),y(1)]);
ddf_dxdy_1 = subs(ddf_dxdy, [X,Y], [x(1),y(1)]);
H = [ddf_ddx_1, ddf_dxdy_1; ddf_dxdy_1, ddf_ddy_1]; % Hessian
S = inv(H); % Search Direction

% Optimization Condition:
while norm(J) > e
    I = [x(i),y(i)]';
    x(i+1) = I(1)-S(1,:)*J';
    y(i+1) = I(2)-S(2,:)*J';
    i = i+1;
    J = [subs(df_dx,[X,Y], [x(i),y(i)]) subs(df_dy, [X,Y],
[x(i),y(i)])]; % Updated Jacobian
    ddf_ddx_1 = subs(ddf_ddx, [X,Y], [x(i),y(i)]);
    ddf_ddy_1 = subs(ddf_ddy, [X,Y], [x(i),y(i)]);
    ddf_dxdy_1 = subs(ddf_dxdy, [X,Y], [x(i),y(i)]);
    H = [ddf_ddx_1, ddf_dxdy_1; ddf_dxdy_1, ddf_ddy_1]; % Updated
Hessian
    S = inv(H); % New Search Direction
end

% Result Table:`
Iter = 1:i;
X_coordinate = x';
Y_coordinate = y';
Iterations = Iter';
T = table(Iterations,X_coordinate,Y_coordinate);

% Plots:
fcontour(f, 'Fill', 'On');
hold on;
plot(x,y,'*-r');
grid on;
```

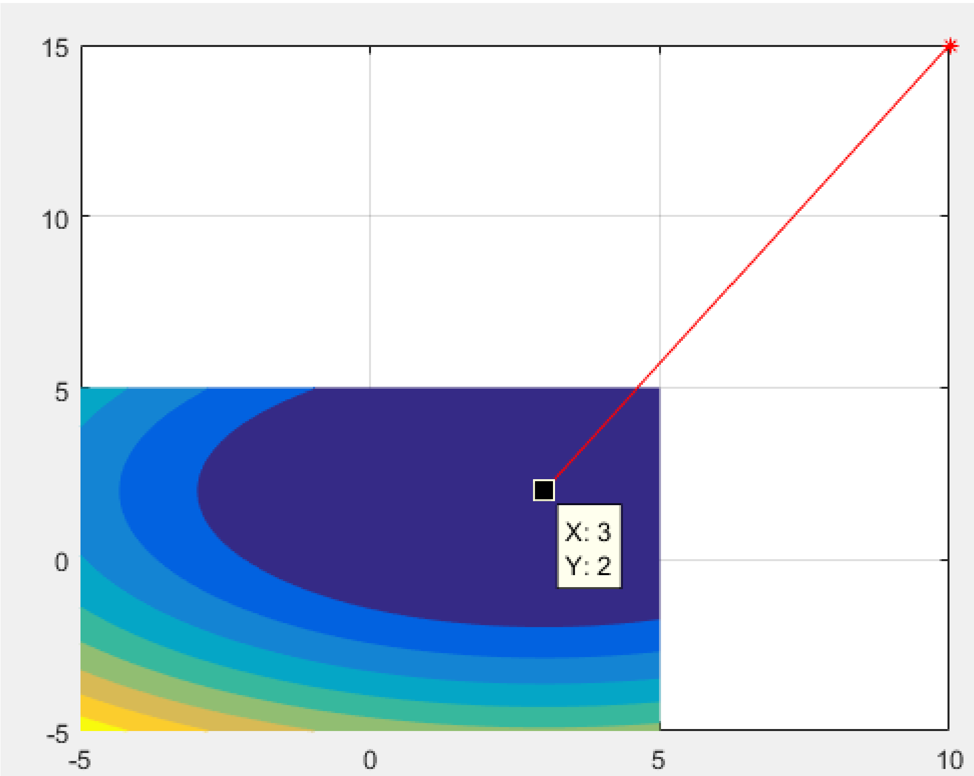
```matlab
% Output:
fprintf('Valor da função objetivo inicial: %d\n\n',subs(f,[X,Y],
[x(1),y(1)]));
if (norm(J) < e)
    fprintf('Mínimo obtido com sucesso...\n\n');
end
fprintf('Número de Iterações para Convergência: %d\n\n', i);
fprintf('Ponto de Mínimo: [%d,%d]\n\n', x(i), y(i));
fprintf('Valor mínimo da função objetivo após a otimização: %f\n\n',
subs(f,[X,Y], [x(i),y(i)]));
disp(T)
```

Output :

```
Command Window

Valor da função objetivo inicial: 112

Mínimo obtido com sucesso...

Número de Iterações para Convergência: 2

Ponto de Mínimo: [3,2]

Valor mínimo da função objetivo após a otimização: 100.000000

    Iterations    X_coordinate    Y_coordinate
    __________    ____________    ____________

    1             10              15
    2              3               2

>>
```

Representação gráfica do método de Newton





Considere o problema de otimização não-linear. (PROBLEMA A). Determine o ponto ótimo X* = (x1, x2) e seu valor de função objetivo f(X*) que minimiza o problema dado através do método de Newton, para X0= [6,6]. (Anexe o desenvolvimento no classroom) *

5 pontos

(PROBLEMA A)

$$\text{Minimizar:} \left(\frac{x_1-3}{6}\right)^2 + \left(\frac{x_2-2}{4}\right)^2 + 100$$

$$x^* = x_0 - \left(F(x_0)\right)^{-1} \nabla f(x_0)$$

$$f(x) = \left(\frac{x_1 - 3}{6}\right)^2 + \left(\frac{x_2 - 2}{4}\right)^2 + 100$$

$$x_0 = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

Substituindo as variáveis x_1 por x e x_2 por y por motivo de facilidade:

$$f(x) = \left(\frac{x - 3}{6}\right)^2 + \left(\frac{y - 2}{4}\right)^2 + 100$$

Gradiente:

$$\nabla f(x, y) = \begin{bmatrix} \frac{x - 3}{18} \\ \frac{y - 2}{8} \end{bmatrix}$$

$$\nabla f(6, 6) = \begin{bmatrix} 0.16 \\ 0.5 \end{bmatrix}$$

Hessiana:

$$F(x) = \begin{bmatrix} \frac{1}{18} & 0 \\ 0 & \frac{1}{8} \end{bmatrix}$$

$$F(x)^{-1} = \begin{bmatrix} 18 & 0 \\ 0 & 8 \end{bmatrix}$$

$$x^* = x_0 - \left(F(x)\right)^{-1} \nabla f(x_0)$$

$$x^* = \begin{bmatrix} 6 \\ 6 \end{bmatrix} - \begin{bmatrix} 18 & 0 \\ 0 & 8 \end{bmatrix} \cdot \begin{bmatrix} 0.16 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 3{,}12 \\ 2 \end{bmatrix}$$

$$x^* = \begin{bmatrix} 3{,}12 \\ 2 \end{bmatrix}$$

$$f(x^*) = \left(\frac{3{,}12 - 3}{6}\right)^2 + \left(\frac{2 - 2}{4}\right)^2 + 100$$

$$f(x^*) = 100{,}0004$$

Determine o ponto ótimo X* = (x1, x2) e seu valor de função objetivo f(X*) que minimiza o "problema A" através do método do Gradiente, dado X0 = [0, 3]. (Anexe o desenvolvimento no classroom) *

5 pontos

(PROBLEMA A)

$$\text{Minimizar:} \left(\frac{x_1 - 3}{6}\right)^2 + \left(\frac{x_2 - 2}{4}\right)^2 + 100$$

$$x_0 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$f(x) = \left(\frac{x_1 - 3}{6}\right)^2 + \left(\frac{x_2 - 2}{4}\right)^2 + 100$$

Substituindo as variáveis x_1 por x e x_2 por y por motivo de facilidade:

$$f(x, y) = \left(\frac{x - 3}{6}\right)^2 + \left(\frac{y - 2}{4}\right)^2 + 100$$

$$\frac{\partial f}{\partial x} = \frac{x - 3}{18}$$

$$\frac{\partial f}{\partial y} = \frac{y - 2}{8}$$

$$\nabla f(x, y) = \begin{bmatrix} \frac{x - 3}{18} \\ \frac{y - 2}{8} \end{bmatrix}$$

$$\nabla f(6, 6) = \begin{bmatrix} 0.16 \\ 0.5 \end{bmatrix}$$

$$\vec{x}^{(i+1)} = \vec{x}^{(i)} - h\nabla\vec{f}$$

$i = 0$

$$1° \, iteração$$

$$\vec{x}^{(1)} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} - h \begin{bmatrix} 0.16 \\ 0.5 \end{bmatrix} = \begin{bmatrix} -0{,}16h \\ 3 - 0{,}5h \end{bmatrix}$$

Substituindo na função:

$$f(x) = \left(\frac{(-0.16h) - 3}{6}\right)^2 + \left(\frac{(3 - 0.5h) - 2}{4}\right)^2 + 100$$

$$f(x) = \frac{2.3524h^2 - 5.16h + 45}{144} + 100$$

$$\frac{\partial f}{\partial h} = 0.032651312h - 0.0358104 = 0$$

$$h = 1.09675$$

$$\vec{x}^{(1)} = \begin{bmatrix} -0{,}16(1.09675) \\ 3 - 0{,}5(1.09675) \end{bmatrix} = \begin{bmatrix} -0.17548 \\ 2.451625 \end{bmatrix}$$

$$x_1 = \begin{bmatrix} -0.17548 \\ 2.451625 \end{bmatrix}$$

$$2° \, iteração$$

$$\nabla f(x, y) = \begin{bmatrix} \dfrac{x - 3}{18} \\ \dfrac{y - 2}{8} \end{bmatrix}$$

$$\nabla f(-0.17548, 2.451625) = \begin{bmatrix} -0.17641 \\ 0.056453125 \end{bmatrix}$$

$$\vec{x}^{(2)} = \begin{bmatrix} -0.17548 \\ 2.451625 \end{bmatrix} - h \begin{bmatrix} -0.17641 \\ 0.056453125 \end{bmatrix} = \begin{bmatrix} -0.17548 + 0.17641h \\ 2.451625 - 0.056453125h \end{bmatrix}$$

Substituindo na função:

$$f(x) = \left(\frac{(-0.17548 + 0.17641h) - 3}{6}\right)^2 + \left(\frac{(2.451625 - 0.056453125h) - 2}{4}\right)^2 + 100$$

$$f(x) = \frac{0.15316\, h^2 - 4.94041h + 42.17037}{144} + 100$$

$$\frac{\partial f}{\partial h} = 0.0021258608h - 0.0342864454 = 0$$

$$h = 16.12826$$

$$\vec{x}^{(2)} = \begin{bmatrix} -0.17548 + 0.17641\,(16.12826) \\ 2.451625 - 0.056453125\,(16.12826) \end{bmatrix} = \begin{bmatrix} 2.66970 \\ 1.54113 \end{bmatrix}$$

$$x_2 = \begin{bmatrix} 2.66970 \\ 1.54113 \end{bmatrix}$$

$3°\ iteração$

$$\nabla f(x,y) = \begin{bmatrix} \frac{x-3}{18} \\ \frac{y-2}{8} \end{bmatrix}$$

$$\nabla f(2.66970, 1.54113\ ) = \begin{bmatrix} -0.01835 \\ -0.05735875 \end{bmatrix}$$

$$\vec{x}^{(3)} = \begin{bmatrix} 2.66970 \\ 1.54113 \end{bmatrix} - h \begin{bmatrix} -0.01835 \\ -0.05735875 \end{bmatrix} = \begin{bmatrix} 2.6697 + 0.01835h \\ 1.54113 + 0.05735875h \end{bmatrix}$$

Substituindo na função:

$$f(x) = \left(\frac{(2.6697 + 0.01835h) - 3}{6}\right)^2 + \left(\frac{(1.54113 + 0.05735875h) - 2}{4}\right)^2 + 100$$

$$f(x) = \frac{0.03095\,h^2 - 0.5222h + 2.33144}{144} + 100$$

$$\frac{\partial f}{\partial h} = 0.000429586h - 0.003624415 = 0$$

$$h = 8.43699$$

$$\vec{x}^{(3)} = \begin{bmatrix} 2.6697 - 0.01835\,(8.43699) \\ 1.54113 - 0.05735875\,(8.43699) \end{bmatrix} = \begin{bmatrix} 2.8245187665 \\ 2.0250652002 \end{bmatrix}$$

$$x_3 = \begin{bmatrix} 2.8245187665 \\ 2.0250652002 \end{bmatrix}$$

$$4° \ iteração$$

$$\nabla f(x,y) = \begin{bmatrix} \frac{x-3}{18} \\ \frac{y-2}{8} \end{bmatrix}$$

$$\nabla f(2.8245187665, 2.0250652002\ ) = \begin{bmatrix} -0.00974 \\ -0.05735875 \end{bmatrix}$$

$$\vec{x}^{(4)} = \begin{bmatrix} 2.8245187665 \\ 2.0250652002 \end{bmatrix} - h \begin{bmatrix} -0.00974 \\ 0.0031325 \end{bmatrix} = \begin{bmatrix} 2.82451 + 0.00974h \\ 2.02506 - 0.0031325h \end{bmatrix}$$

Substituindo na função:

$$f(x) = \left(\frac{(2.82451 + 0.00974h) - 3}{6}\right)^2 + \left(\frac{(2.02506 - 0.0031325h) - 2}{4}\right)^2 + 100$$

$$f(x) = \frac{0.00046h^2 - 0.0150871889h + 0.1288389928}{144} + 100$$

$$\frac{\partial f}{\partial h} = 0.0000063848h - 0.00010 = 0$$

$$h = 15.66219$$

$$\vec{x}^{(4)} = \begin{bmatrix} 2.82451 + 0.00974(15.66219) \\ 2.02506 - 0.0031325(15.66219) \end{bmatrix} = \begin{bmatrix} 2.9770597306 \\ 1.9759981898 \end{bmatrix}$$

$$x_4 = \begin{bmatrix} 2.9770597306 \\ 1.9759981898 \end{bmatrix}$$

$$5° \ iteração$$

$$\nabla f(x,y) = \begin{bmatrix} \frac{x-3}{18} \\ \frac{y-2}{8} \end{bmatrix}$$

$$\nabla f(2.9770597306, 1.9759981898\ ) = \begin{bmatrix} -0.001275 \\ -0.00300125 \end{bmatrix}$$

$$\vec{x}^{(5)} = \begin{bmatrix} 2.9770597306 \\ 1.9759981898 \end{bmatrix} - h \begin{bmatrix} -0.001275 \\ -0.00300125 \end{bmatrix} = \begin{bmatrix} 2.97705 + 0.001275h \\ 1.97599 + 0.00300125h \end{bmatrix}$$

Substituindo na função:

$$f(x) = \left(\frac{(2.97705 - 0.001275h) - 3}{6}\right)^2 + \left(\frac{(-1.97599 + 0.49699875h) - 2}{4}\right)^2 + 100$$

$$f(x) = \frac{0.00008h^2 - 0.00153h + 0.0072951309}{144} + 100$$

$$\frac{\partial f}{\partial h} = 0.0000011104h - 0.0000106182 = 0$$

$$h = 9.5625$$

$$\vec{x}^{(5)} = \begin{bmatrix} 2.97705 + 0.001275(9.5625) \\ 1.97599 + 0.00300125\ (9.5625) \end{bmatrix} = \begin{bmatrix} 2.98924 \\ 2.00468 \end{bmatrix}$$

$$x_5 = \begin{bmatrix} 2.98924 \\ 2.00468 \end{bmatrix}$$

$$6°\ iteração$$

$$\nabla f(x, y) = \begin{bmatrix} \frac{x - 3}{18} \\ \frac{y - 2}{8} \end{bmatrix}$$

$$\nabla f(2.98924, 2.00468) = \begin{bmatrix} -0.00059 \\ 0.000585 \end{bmatrix}$$

$$\vec{x}^{(6)} = \begin{bmatrix} 2.98924 \\ 2.00468 \end{bmatrix} - h \begin{bmatrix} -0.00059 \\ 0.000585 \end{bmatrix} = \begin{bmatrix} 2.98924 + 0.00059h \\ 2.00468 - 0.000585h \end{bmatrix}$$

Substituindo na função:

$$f(x) = \left(\frac{(2.98924 + 0.00059h) - 3}{6}\right)^2 + \left(\frac{(2.00468 - 0.000585h) - 2}{4}\right)^2 + 100$$

$$f(x) = \frac{0.00000447h^2 - 0.0001000676h + 0.000660232}{144} + 100$$

$$\frac{\partial f}{\partial h} = 0.000000062h - 0.0000006945 = 0$$

$$h = 11.2$$

$$\vec{x}^{(6)} = \begin{bmatrix} 2.98924 + 0.00059\ (11.2) \\ 2.00468 - 0.000585\ (11.2) \end{bmatrix} = \begin{bmatrix} 2.99584 \\ 1.998128 \end{bmatrix}$$

$$x_6 = \begin{bmatrix} 2.99584 \\ 1.998128 \end{bmatrix}$$

Portanto, o ponto ótimo que minimiza a função é (2.99584, 1.998128)

$$x^* = (2.99584, 1.998128)$$

$$f(x^*) = \left(\frac{2.99584 - 3}{6}\right)^2 + \left(\frac{1.998128 - 2}{4}\right)^2 + 100$$

$$f(x^*) = 100.00000$$

No desenvolvimento do método de Quasi-Newton calculam-se os coeficientes C(k)DFP e C(k)BFGS que representam os coeficientes da família Broyden. Explique qual o papel destes coeficientes no método e represente a equação que demonstra como se dá a sua participação no processo de correção da H(k). * 5 pontos

São coeficientes de correção da inversa da hessiana.

$$H_{k+1} = H_k + C_k(\alpha)$$

Sendo α um número aleatório ∈ [0,1] que irá nos dizer qual dos dois coeficientes da família de Broyden deve ser usado, DFP ou BHGS.

Se α sorteado for igual a 0, será usado o coeficiente de correção de DFP.

Se α sorteado for igual a 1, será usado o coeficiente de correção de BHGS.

Se α estiver entre 0 e 1, o coeficiente será uma combinação do coeficiente de DFP e BHGS.

A cada iteração, a hessiana da próxima iteração será a hessiana atual somada com o coeficiente $C_k(\alpha)$ calculado.

Assim, é feita uma correção da hessiana na iteração, e a cada iteração são feitas novas correções até convergir para uma solução ótima.

Cite ao menos 2 critérios de parada e convergência que possam ser utilizados para a solução ótima de um problema de otimização não-linear, descrevendo o funcionamento de 1 dos métodos. * 4 pontos

Anulação do vetor gradiente: quanto mais próximo o vetor gradiente estiver de zero, mais próximo se está do ponto crítico.

É calculado o módulo dos gradientes da função objetivo, verifica-se se esse valor do gradiente é menor que 0,1% do valor do maior gradiente calculado (geralmente é o valor em módulo do primeiro gradiente calculado na iteração 1, e esse valor do gradiente ficará fixo). A cada iteração é feita uma verificação se o gradiente é menor que 0,1% do maior valor do gradiente fixo, e esse loop será feito até que a condição seja atendida. Então, o valor do gradiente será tão pequeno que se pode considera-lo como zero, e assim convergindo para a solução ótima.

Estabilização do vetor de variáveis de otimização: consiste em verificar o valor das variáveis a cada iteração. Se esses valores tiverem apenas uma pequena variância em relação aos anteriores, então o algoritmo pode parar, pois o erro entre os valores das variáveis é tão pequeno que se entende que aquele ponto é ponto de ótimo da solução

Cite 2 características do método de busca aleatória e justifique se promovem eficiência ou ineficiência quanto à utilização do método. * 4 pontos

É um método de força “bruta”, que faz busca mudando o sentido e a direção sem nem um tipo de critério, muda de forma aleatória, isso faz com que o método não seja tão eficiente, pois essa direção aleatória prejudica na convergência do método.

Ele é um método que consegue apenas uma aproximação da solução ótima, ou seja, a convergência na solução do ponto mínimo x* não é satisfatória. Isso faz com que ele não seja um bom método pra ser utilizado para busca de uma solução ótima.

Qual a principal proposição da utilização do método de Newton Modificado quanto à inviabilidade da utilização do método de Newton para problemas não quadráticos? * 3 pontos

Em um problema não quadrático a inversa da hessiana não conseguirá ter uma boa convergência devido à mudança na curvatura da função.

A proposição para esse tipo de problema não quadrático, é que se faça uma correção a cada iteração.

Usando a fórmula inicial dos métodos de direção de busca:

$$x_{k+1} = x_k + \alpha_k \,.\, d_k$$

E a fórmula de Newton:

$$x^* = x_0 - \left(F(x)\right)^{-1} \nabla f(x_0)$$

A parte $-\left(F(x)\right)^{-1} \nabla f(x_0)$ do método de Newton é definida como uma direção d_k

O α_k são pequenas correções feitas no método a cada iteração para que ele consiga convergir para uma solução ótima.