UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELETROELETRÔNICA
OTIMIZAÇÃO NÃO-LINEAR

TRABALHO PRATICO 1

Aluno: Rafael Lázaro Monteiro Matrícula: 2017435036
Curso: Engenharia de Sistemas

Sumário

1. INTRODUÇÃO 1
2. CARACTERISTICAS DOS MÉTODOS IMPLEMENTADOS. 2
2.1. Método do Gradiente: 2
2.2. Newton Modificado 2
2.3. Método de Broyden 3
2.4. Observação geral sobre os métodos de Newton. 3
2.5. Método do Gradiente Conjugado 4
3. IMPLEMENTAÇÃO 5
3.1. Observações sobre a Implementação 5
4. RESULTADOS 6
5. CONCLUSÃO 10
6. REFERÊNCIA BIBLIOGRÁFICA 11

Lista de Figuras

Figura 1: Trajetória do Algoritmo do Gradiente durante a Otimização de objfun1. 7
Figura 2: Trajetória do Algoritmo e Newton Modificado durante a Otimização de objfun1..... 7
Figura 3: Trajetória do Algoritmo QuaseNewton durante a Otimização de objfun1 7
Figura 4: Trajetória do Algoritmo do Gradiente Conjugado durante a Otimização de objfun1 8
Figura 5: Gráfico que relaciona o número de avaliações da função com o valor da função. 9

Lista de Tabela

Tabela 1: Resultados Obtidos do Processo de Otimização da Função Objfun1. 6
Tabela 2: Valores obtidos a partir da aplicação dos algoritmos de Otimização para a função objfun2. 8
Tabela 3: Ponto retornado pelos algoritmos para objfun2 8

1. INTRODUÇÃO

Os primeiros métodos de otimização (minimização) de funcionais não-lineares foram desenvolvidos a partir da ideia básica de fazer o algoritmo evoluir encontrando novos pontos situados em direções para a função decresça, em relação ao ponto corrente.

O algoritmo mais antigo de otimização se baseia no gradiente da função daí vindo o seu nome de Algoritmo do Gradiente, este busca pontos sobre a reta definida por um ponto e pelo gradiente da função objetivo.

Esse algoritmo toma os valores de pontos menores que o valor do ponto corrente até que se atinja os critérios de convergência.

Os algoritmos de otimização foram aperfeiçoados para melhorar seu desempenho e resposta e a partir do método do gradiente foram feitas varia mudança e daí nasceram outras famílias de algoritmo de otimização. Estes métodos se baseiam em uma direção de busca e assim possuem em comum as seguintes características:

Cada novo ponto e obtido de um processo de otimização unidimensional que tem como ponto de partida o ponto anterior.

A direção na qual e feita a busca unidimensional e uma função das avaliações anteriores da função objetivo.

Tendo isso em mente o objetivo deste trabalho é a implementação de Algoritmos de otimização para realizar a otimização de duas funções de interesse. Os algoritmos implementados nesse trabalho foram:

- Método do Gradiente => GradMethod;
- Método de Newton Modificado => newtonmethod;
- Método Quase Newton => quasenewton;
- Método do Gradiente Conjugado (Fletcher-Reeves) => conjgrad;

As funções estão implementadas como:

- Objfun1
- Objfun2

2. CARACTERISTICAS DOS MÉTODOS IMPLEMENTADOS.

2.1. Método do Gradiente:

Inserido na família dos métodos de direção de busca, o método gradiente consiste em determinar a direção de busca para o ponto de ótimo de uma função através do gradiente da função em determinados pontos. Ou seja, dado um determinado ponto inicial, o primeiro passo do algoritmo é determinar o gradiente da função neste ponto, de tal maneira que a direção de busca será de valor oposto ao gradiente calculado, gerando uma nova solução ao longo desta direção de busca.

Abaixo são apresentadas algumas características do método;

- O critério de redução do gradiente abaixo de um determinado delta não é na maioria das vezes apropriado, na pratica são utilizados critérios mais realistas.
- Algoritmo de maior descida a cada iteração minimiza-se f ao longo da direção de maior descida, a do gradiente.
- O algoritmo é globalmente convergente, converge para um mínimo local de f a partir de qualquer ponto x_0 inicial.

2.2. Newton Modificado

O método de Newton se enquadra na família das aproximações quadráticas, métodos que utilizam aproximações da função em termos da série de Taylor em que, a partir das condições de primeira ordem se obtém, através do gradiente e da Hessiana, uma estimativa para o ponto de ótimo da função objetivo, o que tende a convergir mais rapidamente que o método gradiente. Caso a função não seja quadrática, foi criado o método de Newton modificado, de tal forma que o algoritmo garanta que haverá uma diminuição monotonica do valor da função objetivo. Ao desenvolvermos a nova solução a ser iterada, o valor de “alpha” no algoritmo ´e otimizado para sempre buscarmos o seu menor valor possível.

Abaixo são apresentadas algumas características do algoritmo

- Geralmente, espera-se obter uma convergência quadrática, desde que um valor inicial próximo da solução seja conhecido.
- Com a modificação, o método de Newton se torna globalmente convergente, com convergência quadrática na vizinhança de x.

- Se o delta escolhido for muito pequeno a variação da função se torna próxima de um singular, se o delta for muito grande compromete-se a ordem de convergência quadrática.

2.3.Método de Broyden

Pertencente aos métodos Quase-Newton, o BFGS trata se de um método que limita o uso de memória, capaz de lidar com problemas com várias variáveis e de alta complexidade. O algoritmo consiste em cada iteração buscar uma aproximação para a matriz Hessiana a fim de obter sua inversa.

- O método de Broyden é uma generalização do método da Secante para sistemas de equações não lineares, sendo pertencente a uma classe de métodos chamados quase-Newton. Uma vantagem destes métodos é que estes substituem a matriz Jacobiana J (x_k) do método de Newton por uma matriz de aproximação B_k que se atualiza em cada iteração.
- A desvantagem dos métodos quase-Newton é a perda da convergência quadrática do método de Newton, sendo substituída, em geral, por uma convergência chamada super linear.
- Outra desvantagem dos métodos quase-Newton é que, ao contrário do método de Newton, esses não são auto corretores. O método de Newton geralmente corrigirá o erro de arredondamento ao decorrer das iterações, não sendo assim no método de Broyden, a menos que se incorpore medidas especiais de correção. O método de Broyden é descrito pelo algoritmo abaixo. Para a iteração inicial, a matriz B será a matriz Jacobiana do sistema para a estimativa inicial x0.

2.4.Observação geral sobre os métodos de Newton.

A vantagem dos métodos de Newton e quase-Newton na resolução de sistemas de equações não lineares é a sua rapidez de convergência, uma vez que se conhece um valor inicial suficientemente próximo da solução exata. Assim, uma de suas debilidades consiste em requererem uma boa aproximação inicial da solução para garantir a convergência.

2.5.Método do Gradiente Conjugado

O Método dos Gradientes Conjugados (CGM) é um algoritmo para encontrar o mínimo local de uma função de n variáveis, supondo que o gradiente da função possa ser calculado. Ele usa direções conjugadas (ortogonais) ao invés do gradiente local para buscar o mínimo. Se a vizinhança de um mínimo tem o aspecto de um vale suave, o mínimo é alcançado em poucos passos, de maneira mais rápida do que usando o Método da Máxima Descida (Steepest descent). O CGM é um método efetivo para a resolução de sistemas matriciais com matrizes simétricas definidas positivas. Ele é o mais antigo e o mais conhecido método não estacionário. O método caminha gerando vetores a cada iteração. Os resíduos r (i) correspondendo às iterações e as direções de busca p (i) são utilizados para atualizar a iteração seguinte. A cada iteração do método, produtos internos são calculados a fim de atualizar os escalares que são definidos para satisfazer certas condições de ortogonalidade. Isso pode ser interpretado como a busca pela mínima energia E de um sistema linear A.x = b. A energia do sistema é mínima quando o resíduo r = b - A.x se anula.

3. IMPLEMENTAÇÃO

Os Algoritmos implementados para este trabalho apresentam algumas peculiaridades, estas são:

- As condições de parada escolhidas para implementação foram, um número de avaliação da função de no máximo 2000 (ncf_{max} = 2000).
- Redução da variação do valor da função entre duas interações para um valor abaixo de uma determinada tolerância.

Além das considerações em relação ao critério de parada, o Método Quase Newton foi implementado considerando a influência dos algoritmos DFP e BFGS, a escolha do efeito de cada sobre o Algoritmo é escolhido pelo usuário que pode entrar com a variável “ETA” que pode variar de 0 a 1, que fará com que haja a influência só de um ou de outro (se ETA =1, só o BFGS influencia se 0 só o DFP), o padrão para execução do trabalho e avaliação neste documento foi de um ETA = 0.5.

Os demais valores usados para a simulação são:

- `x1= [-1.2 ; 1.0];`
- `x2 = [ 5.0 ;5.0 ;5.0 ;5.0 ;5.0 ;5.0 ;5.0 ;5.0 ;5.0 ;5.0];`
- `tol = 0.0001;`
- `delta= 0.0001;`
- `NFC = 2000;`
- `s = 0.001;`
- `Epsi = 0.0001;`
- `Eta = 0.5;`

Os valores acima são os seguintes, os dois primeiros valores são os pontos x1 para a função objfun1, e x2 para objfun2, tol e o valor mínimo de variação da função para um valor menor que tol de variação da função o algoritmo é interrompido, delta e o valor para análise de diferença finita do gradiente, NFC é o número de avaliações de função s e Epsi são coeficientes de precisão.

3.1. Observações sobre a Implementação

Os algoritmos todos foram implementados com uma pequena explicação em relação as entradas, a função dos mesmos e os valores retornados, essa explicação está em inglês para todos os algoritmos o que é uma pratica comum tendo que o inglês e a linguagem básica utilizada em programação.

4. RESULTADOS

As seguir são apresentados os resultados dos processos de otimização aplicados para as seguintes funções:

Objfun1:

$$f_1(\mathbf{x}) = \sum_{i=1}^{n-1} \left[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\right]$$

Objfun2:

$$f_2(\mathbf{x}) = \sum_{i=1}^{n} \left[\sum_{j=1}^{i} (x_j - j)^2\right]$$

Como resultado da aplicação dos algoritmos a função objfun1 tem-se a seguinte tabela, que mostra os valores observados durante o processo de otimização para a função objfun1.

Tabela 1: Resultados Obtidos do Processo de Otimização da Função Objfun1.

Método	Ponto de mínimo (x1,x2)		Número de Avaliações da função	Número de Iterações	Tempo de Execução (s)
Método do Gradiente	0,84293	0,70808	2023	104	0,01589
Newton Modificado	0,99061	0,98079	2005	131	0,02114
Quase Newton	-1,02667	1,06189	2004	182	0,01275
Gradiente Conjugado	1,00088	1,00232	606	55	0,00424

Com base nos métodos implementados foram plotados gráficos que mostram a evolução da trajetória do algoritmo durante o processo de otimização. Os gráficos foram todos plotados com a mesma configuração, os pontos foram configurados para serem ressaltados pontos de variando de 5 em 5. Logo a comparação entre os mesmos é válida.

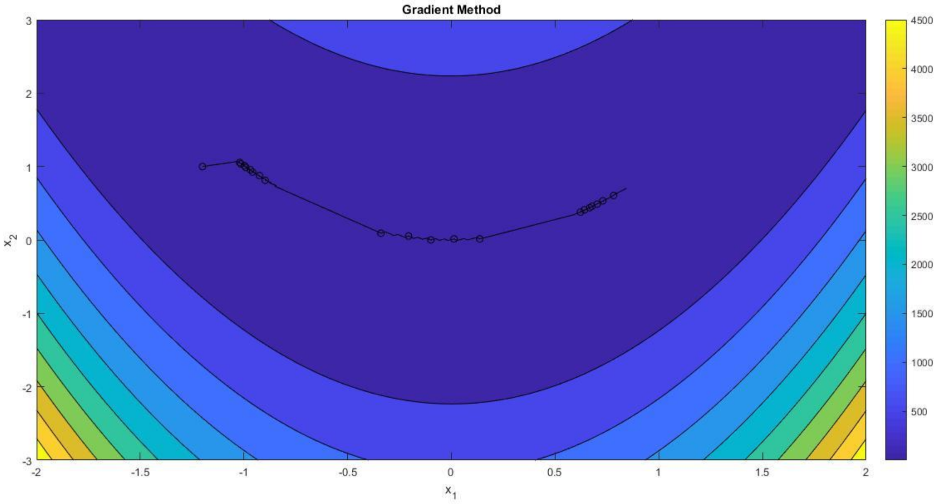

Figura 1: Trajetória do Algoritmo do Gradiente durante a Otimização de objfun1.

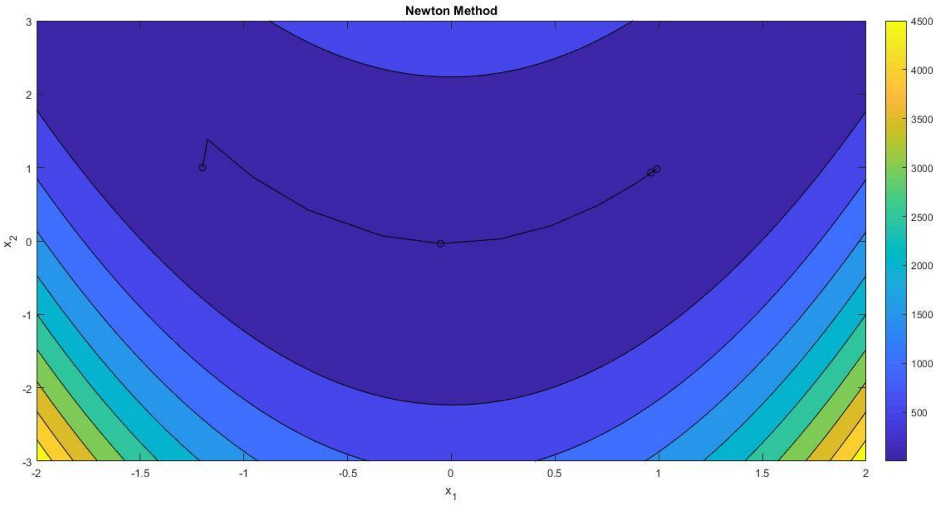

Figura 2: Trajetória do Algoritmo e Newton Modificado durante a Otimização de objfun1

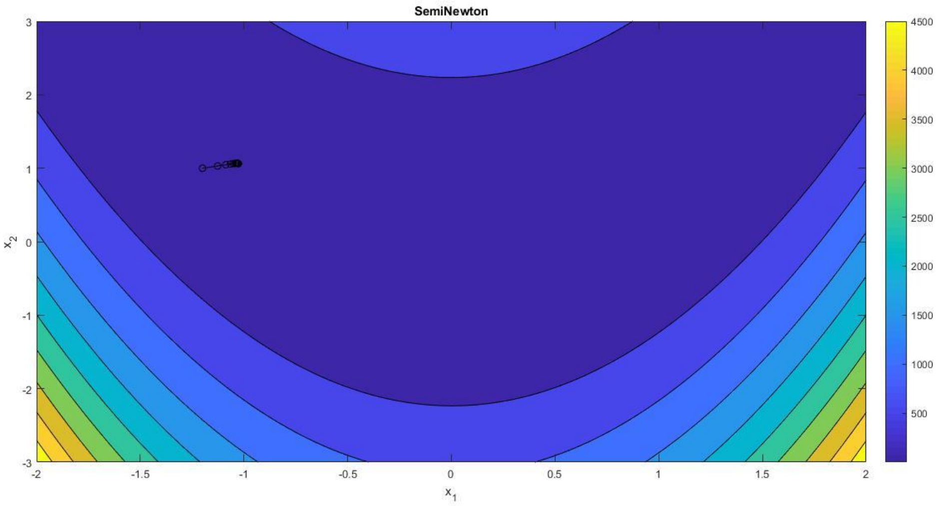

Figura 3: Trajetória do Algoritmo QuaseNewton durante a Otimização de objfun1

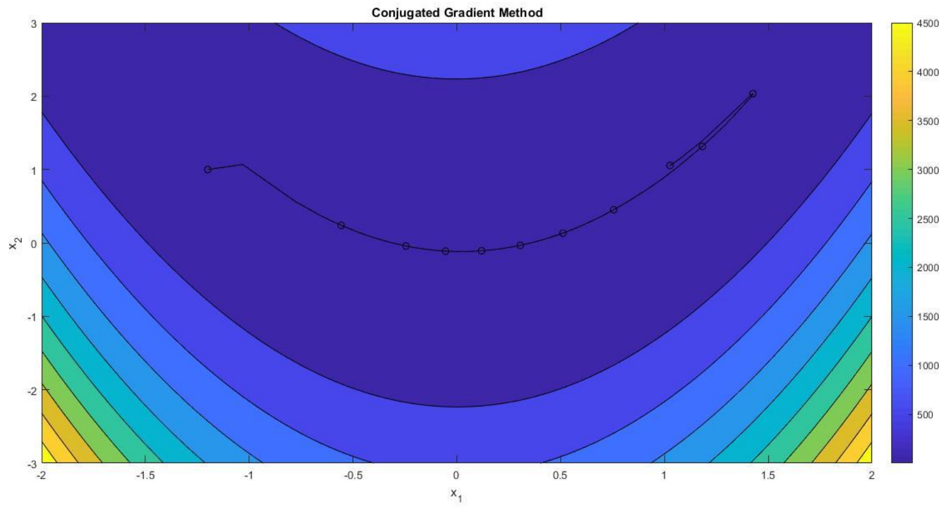

Figura 4: Trajetória do Algoritmo do Gradiente Conjugado durante a Otimização de objfun1

Como resultado da aplicação dos algoritmos a função objfun2 tem-se a seguinte tabela com os valores obtidos e observados durante a otimização da função usando os algoritmos implementados.

Tabela 2: Valores obtidos a partir da aplicação dos algoritmos de Otimização para a função objfun2.

Método	Número de Avaliações da função	Número de Iterações	Tempo de Execução
Método do Gradiente	1046	28	0,00429
Newton Modificado	66	1	0,00207
Quase Newton	2005	182	0,01476
Gradiente Conjugado	260	9	0,00093

Os pontos de mínimo encontrados para a função para cada algoritmo são mostrados na tabela abaixo.

Tabela 3: Ponto retornado pelos algoritmos para objfun2

Ponto de Mínimo Calculado pelos Algoritmos para objfun2										
Método do Gradiente	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
Método do Gradiente	1,001365	1,99995	2,99995	3,99995	4,99995	5,99995	6,99995	7,99995	8,999949	9,993029
Newton Modificado	0,99993	1,999934	2,999942	3,999945	4,99995	5,999958	6,999958	7,999972	8,999965	9,999986
Quase Newton	4,638459	4,754828	4,854038	4,935843	4,999997	5,04625	5,074347	5,084032	5,075045	5,047123
Gradiente Conjugado	1,000192	1,999933	2,999963	3,999944	4,99995	5,999965	6,99993	7,999954	8,999963	9,999992

Considerando os dados obtidos na simulação foi plotado o gráfico abaixo que relaciona o valor da função no ponto e número de avaliações da função realizadas até a iteração corrente.

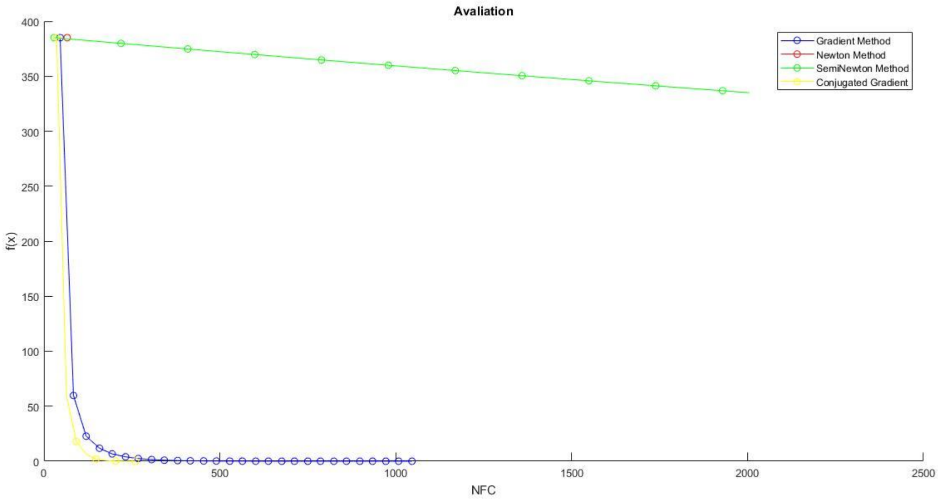

Figura 5: Gráfico que relaciona o número de avaliações da função com o valor da função.

5. CONCLUSÃO

Com o resultado da observação dos dados apresentados pode se notar que em primeiro lugar, para função “objfun1” os algoritmos não convergiram para o ponto de mínimo da função dentro dos critérios de adotados para avaliação (o que faz sentido pois a função em questão é uma variante da função de Rosenbrock, a qual é uma função difícil para que haja conversão para o ponto de mínimo). Também pode-se notar pela observação dos gráficos que o algoritmo Quase Newton apresentou passos muitos pequenos mantendo-se próximo ao mesmo ponto durante o processo de otimização com os parâmetros passados. Enquanto os outros algoritmos apresentaram passos maiores, somado a isso o algoritmo com menor número de iterações foi o do Gradiente Conjugado.

Com relação a função objfun2 pode-se observar que apenas o Algoritmo de Quase Newton não convergiu, sendo que o Newton Modificado apresentando a maior eficiência tanto com tempo como número iterações, resultado que também pode ser observado no gráfico ncf x f(x).

6. REFERÊNCIA BIBLIOGRÁFICA

1. Otimização Escalar e Vetorial Volume 2: Otimização Escalar .Professor: Ricardo H. C. Takahashi Belo Horizonte, Janeiro de 2007