

Relatório 5

Método de Descida de Gradiente

Algoritmo de Optimização

Cristiano Lopes Moreira

Matrícula: 119103-0

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
13/08/2019	1	2º. Semestre de 2019		PEL_216_Relatório_5_Cristiano_Moreira.doc			1 (11)

Sumário

1.	Introdução	3
2.	Desenvolvimento teórico	3
2.1.	Descrição do problema:	4
2.2.	Algoritmo de Descida de Gradiente:	5
3.	Proposta de implementação	6
4.	Experimentação e Resultados	9
5.	Trabalhos Correlatos	10
6.	Conclusão	11
7.	Referências bibliográficas	11

Aluno		RA/Matrícula	Professor	Tipo	
Cristiano Lopes Moreira		119103-0	Dr Reinaldo Bianchi	Relatório de implementação	
Data	Versão	Turma	Nome do arquivo		Página
13/08/2019	1	2º. Semestre de 2019	PEL_216_Relatório_5_Cristiano_Moreira.doc		2 (11)

1. Introdução

Para problemas em que é necessário encontrar o ponto de mínimo no comportamento de uma função, é possível utilizar mecanismos computacionais para realizar a varredura de todas as possíveis entradas dessa função, porém, as funções que contêm muitas possibilidades geram buscas longas ou até infinitas, o que torna essa estratégia inviável para esses casos.

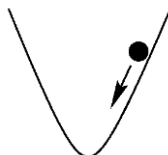
Sistemas computacionais desejáveis devem realizar suas rotinas de forma rápida, com baixo custo, e baixo uso de memória; nos casos de validações de funções, devem avaliar $f(x)$ na menor quantidade de vezes possíveis.

A metodologia de descida de gradiente utiliza a taxa de variação de uma função, por sua derivada para inferir aproximações sobre os pontos de mínimo de uma determinada função, de forma a evitar a necessidade de buscar todas possibilidades desta função.

Este trabalho tem o objetivo de implementar e validar o algoritmo de otimização de descida de gradiente para encontrar o valor mínimo local de uma função.

2. Desenvolvimento teórico

Na definição matemática, gradiente, representado pelo símbolo ∇ , é um vetor, a partir de um ponto em uma função, que indica a direção (sentido) na qual existe o maior valor de deslocamento (variação) da função. Quando a variação de uma função em um determinado ponto é igual a zero, ou seja, seu gradiente é zero, essa função se encontra em um ponto de inflexão, ponto de máximo ou mínimo local, ponto estacionário da função $f(x)$.



Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
13/08/2019	1	2º. Semestre de 2019		PEL_216_Relatório_5_Cristiano_Moreira.doc			3 (11)

Quando é possível calcular, em um determinado ponto x , não somente a função $f(x)$ mas também o seu gradiente (vetor de sua derivada) $\nabla f(x)$. É possível encontrar o mínimo local desta função através de passos decrescentes em direção ao gradiente de descida. Inicia-se no ponto inicial (x_0) , e quantas vezes forem necessárias, move-se do ponto (x_i) ao ponto (x_{i+1}) através da minimização do ponto x , no eixo x , na direção decrescente do gradiente $-\nabla f(x_i)$. Esse método de descida de gradiente irá realizar vários pequenos passos, “descendo o vale”, até encontrar o ponto de menor valor de $f(x_i)$. Portanto, é natural que, nos passos de busca do mínimo de uma função caminhemos na direção contrária ao gradiente, isto é, para um dado ponto x_i temos:

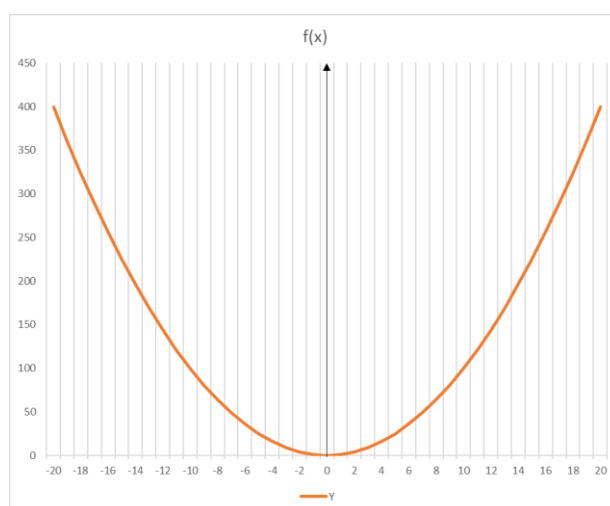
$$x_{i+1} = x_i - \nabla f(x_i)$$

2.1. Descrição do problema:

Deseja-se encontrar o mínimo das funções abaixo utilizando o método de descida de gradiente, fazendo variação na taxa de aprendizagem.

Primeira função: x^2 , sendo o ponto $x_0=2$

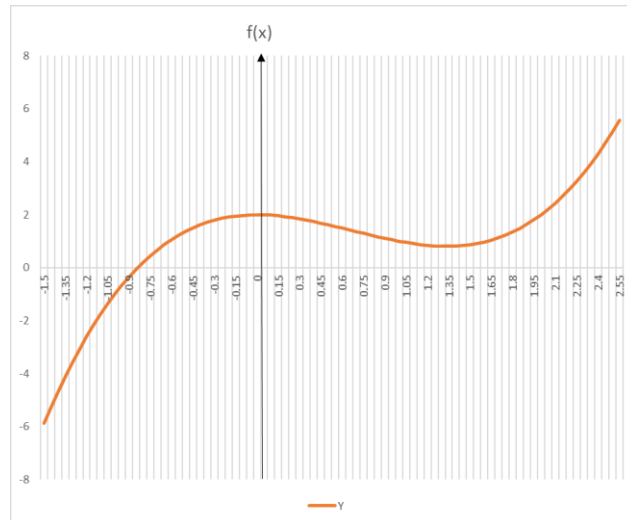
Derivada $2x$



Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 13/08/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_216_Relatório_5_Cristiano_Moreira.doc		Página 4 (11)

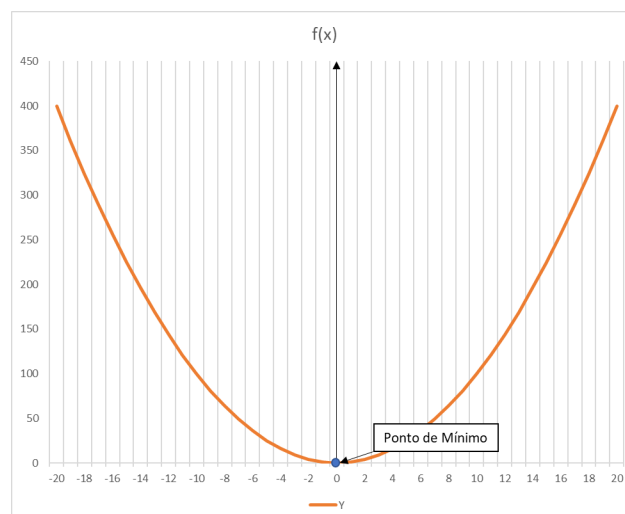
Segunda função: $x^3 - 2x^2 + 2$, sendo o ponto $x_0=2$

Derivada: $3x^2 - 4x$



2.2. Algoritmo de Descida de Gradiente:

O algoritmo de descida de gradiente utiliza de uma metodologia de otimização via derivadas de primeira ordem para encontrar o mínimo de uma função. Suas operações consistem em subtrair o gradiente do valor do ponto (x) da amostragem até que seja encontrado o ponto do mínimo local, quando a derivada da função no ponto for zero, ou próxima a um fator de erro definido no algoritmo.



Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
13/08/2019	1	2º. Semestre de 2019		PEL_216_Relatório_5_Cristiano_Moreira.doc			5 (11)

Seus métodos principais são:

construtor (txErro, txAprendizagem, fx): Recebe a taxa de erro e aprendizagem, e a função que se deseja encontrar o ponto de mínimo.

calc (x0): realiza as interações para se aproximar do ponto objetivo.

vInflexao (): retorna o ponto x em que a função tem um valor mínimo local

3. Proposta de implementação

É proposta a implementação de 2 objetos, o objeto GRADIENTE, objeto base que recebe a função a qual se deseja encontrar o mínimo local; e objeto Funcd, objeto que transporta as informações da função.

Estrutura do objeto gradiente

O Algoritmo gradiente será implementado via Classe, e inicializado pelo método de construção que recebe as taxas de erro e aprendizagem, e a função a qual se deseja calcular o mínimo. O gradiente utiliza o método 'calc', responsável por receber o ponto 'x' inicial e ir aproximando, pela metodologia de gradiente, do ponto de mínimo da função; qTpassos, responsável por quantificar quantos passos foram necessários para alcançar o objetivo; vInflexao, responsável por mostrar o valor de 'x' no ponto de mínimo da função e vIFxInflexao, responsável por mostrar o valor do ponto do 'y', (fx), da função. Além dos métodos o objeto gradiente armazena as informações da taxa de erro, e taxa de aprendizado.

Será utilizado um objeto base Funcd, que contém a função e a derivada da função a qual se pretende encontrar o mínimo local

Funcd
<<constructor>> Funcd() <<destructor>> ~Funcd() + funcao(double x) : double + derivada(double x): double

Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 13/08/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_216_Relatório_5_Cristiano_Moreira.doc		Página 6 (11)

O Objeto gradiente

gradiente
<ul style="list-style-type: none"> - passos: integer - erro: double - aprendizado: double - vInflexao: double - vIFxInflexao: double - pointFx: class Functd
<pre> <<constructor>> gradiente(Inicia: int) <<destructor>> gradiente (Inicia: int) + calc (double x0) + qTpassos () + inflexao () + funcFx () </pre>

Pseudocódigo:

```

calc(x0)
    x=x0;
    vFx = funcao(x0)
    dFx = derivada(x0)
    enquanto (dFx != 0)
        dFx = derivada(x)
        x=x- txAprendizado*dFx
        dFx = derivada(x)

    vInflexao=x
    vIFxInflexao= funcao

inflexao()
    retornar vInflexao
    retorna verdadeiro

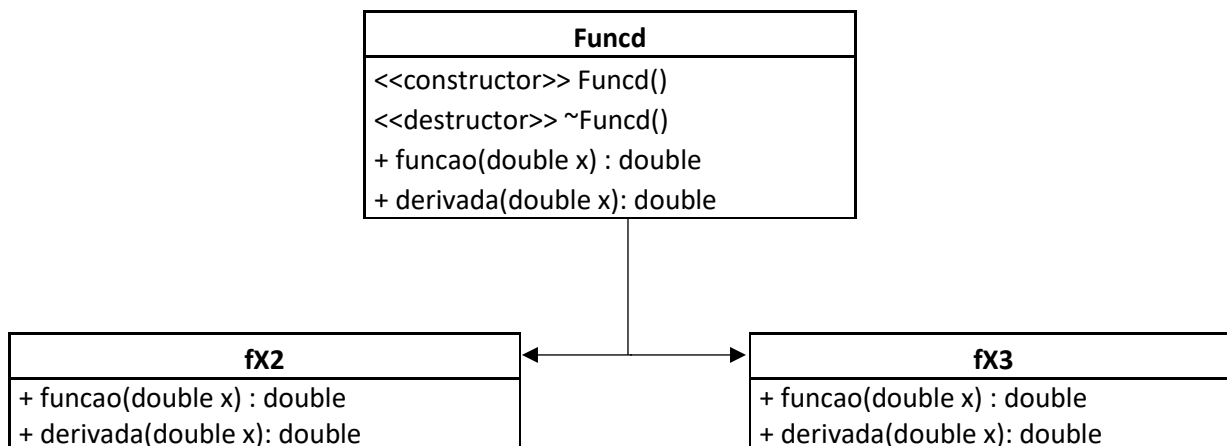
funcfx()
    retorna vIFxinflexao

```

Aluno		RA/Matrícula	Professor	Tipo	
Cristiano Lopes Moreira		119103-0	Dr Reinaldo Bianchi	Relatório de implementação	
Data	Versão	Turma	Nome do arquivo		Página
13/08/2019	1	2º. Semestre de 2019	PEL_216_Relatório_5_Cristiano_Moreira.doc		7 (11)

Herança da estrutura de função para formação de fX2 e fX3

Os Algoritmos fX2 e fX3 serão implementados via classe com herança da Classe Funcd, com métodos 'funcao' e 'derivada' implementados pela utilização de polimorfismo.



Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
13/08/2019	1	2º. Semestre de 2019		PEL_216_Relatório_5_Cristiano_Moreira.doc			8 (11)

4. Experimentação e Resultados

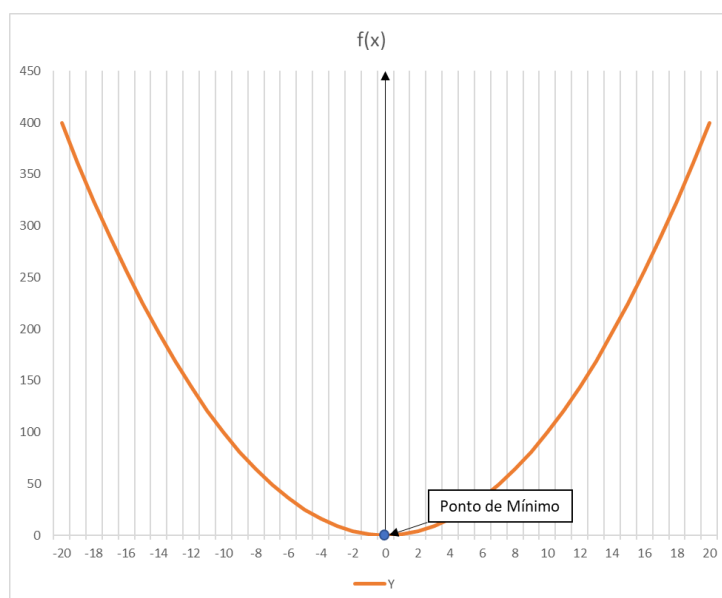
Foram gerados ensaios utilizando o método de descida de gradiente com o valor da taxa de aprendizagem variando entre 0.1 a 1.0 e observado o número de passos necessário, para cada taxa de aprendizagem, para encontrar o objetivo final, valor mínimo da função.

Problema 01: $f(x) = x^2$

Taxa de erro (0.0001)

Taxa de aprendizado (0.1)

$X_0=2$



Aprendizado	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Passos	79	35	20	11	1	11	20	35	79	∞
X	0	0	0	0	0	0	0	0	0	-2/2
y	0	0	0	0	0	0	0	0	0	4

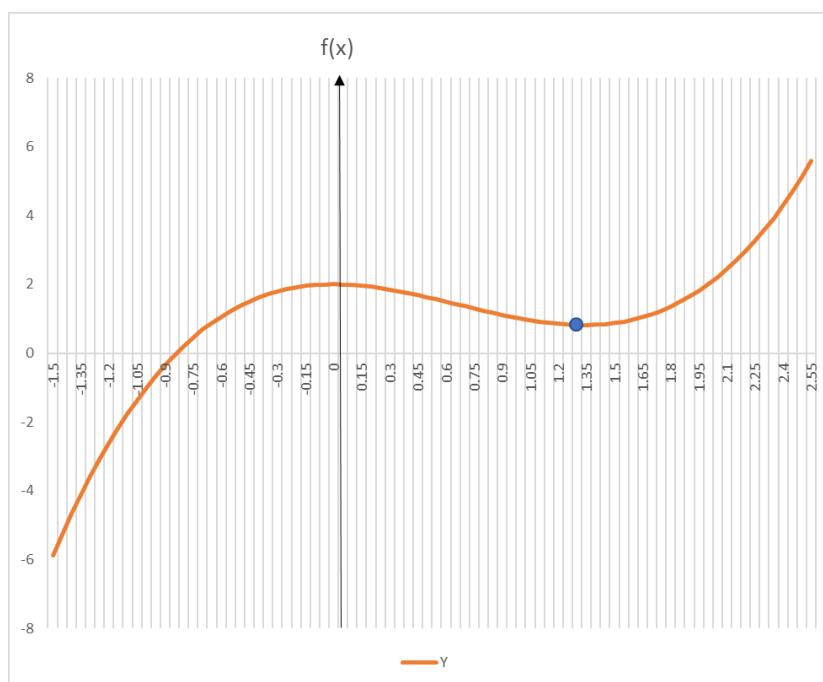
Aluno Cristiano Lopes Moreira		RA/Matrícula 119103-0	Professor Dr Reinaldo Bianchi	Tipo Relatório de implementação	
Data 13/08/2019	Versão 1	Turma 2º. Semestre de 2019	Nome do arquivo PEL_216_Relatório_5_Cristiano_Moreira.doc		Página 9 (11)

Problema 02: $f(x) = x^3 - 2x^2 + 2$

Taxa de erro (0.0001)

Taxa de aprendizado (0.1)

$X_0=2$



Aprendizado	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Passos	33	11	12	28	1	∞	∞	∞	∞	∞
X	1.33333	1.33333	1.33333	1.33333	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
y	0.81481	0.81481	0.81481	0.81481	2	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$

5. Trabalhos Correlatos

NECULAI A, **A Dai-Liao conjugate gradient algorithm with clustering of eigenvalues**, Numerical Algorithms, April 2018, Volume 77, Issue 4, p 1273–1282

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
13/08/2019	1	2º. Semestre de 2019		PEL_216_Relatório_5_Cristiano_Moreira.doc			10 (11)

6. Conclusão

Os resultados mostram para a função $f(x)=x^2$, o algoritmo de descida de gradiente se mostra bem eficiente para encontrar o mínimo local; as variações na taxa de aprendizado mostraram variações pouco significantes na quantidade de passos para encontrar a solução, com a exceção da taxa de aprendizado 1.0 que leva a busca pelo mínimo a um loop pelo fato do gradiente levar a função para os pontos -2 e 2 constantemente.

Para a função cúbica em $f(x) = x^3 - 2x^2 + 2$, o algoritmo encontrou o mínimo local, via metodologia de descida do gradiente, quando a taxa de aprendizado estava entre 0.1 e 0.4, que movimentou a referência 'x' antes de ultrapassar o ponto de inflexão da função e obteve o resultado apresentando o mínimo local. Para os casos em que a taxa de aprendizado foi superior a 0.5, a referência ultrapassou o ponto de inflexão da função e o algoritmo começou a descer até o infinito, não apresentando um ponto de parada.

7. Referências bibliográficas

W. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, B. P. FLANNERY, **Numerical Recipes – The Art of Scientific Computing**, 3rd ed. Cambridge University Press, 2007.

Aluno		RA/Matrícula		Professor		Tipo	
Cristiano Lopes Moreira		119103-0		Dr Reinaldo Bianchi		Relatório de implementação	
Data	Versão	Turma		Nome do arquivo			Página
13/08/2019	1	2º. Semestre de 2019		PEL_216_Relatório_5_Cristiano_Moreira.doc			11 (11)