Praca domowa nr 1

Napisz kod w R i dodaj własne komentarze wyjaśniające jego działanie.

1. Liczby Fibonacciego (kolejna liczba jest sumą dwóch poprzednich) Napisz funkcję która obliczy wartość n-tej liczby Fibonacciego
2. Liczby pierwsze Napisz funkcję która wypisuje wszystkie liczby pierwsze w zakresie od 0 do n

In [1]:

```
fibo<-function(n){#nth Fibonacci number
    varhelp1=1
    varhelp2=1
    varhelp3=0
    if(n==1||n==2)
      return(1)
    for(i in 3:n){
      varhelp3=varhelp1+varhelp2
      varhelp1=varhelp2
      varhelp2=varhelp3
    }
return(varhelp3)
}
```

In [2]:

```
fibo(4)
```

3

In [3]:

```
primes <- function(n){
    result<-c()
    for (checked_number in 2:n){
        no_divisors<-0
        for(checked_divisor in 1:n){
            if(checked_number%%checked_divisor==0){
                no_divisors<-no_divisors+1
            }
        }
        if(no_divisors==2){
            result<-c(result,checked_number)
        }
    }
    return(result)
}
```

In [4]:

```
primes(100)
```

2 · 3 · 5 · 7 · 11 · 13 · 17 · 19 · 23 · 29 · 31 · 37 · 41 · 43 · 47 · 53 · 59 · 61 · 67 · 71 · 73 · 79 · 83 · 89 · 97

Praca domowa nr 2

2.1 Optymalizacja – ograniczenia równościowe. Znajdź maksimum i minimum funkcji.
$f(x, y) = 2x - y$ pod warunkiem
$g(x, y) = x^2 + y^2 - 5 = 0$

2.1

$\mathcal{L} = 2x - y - \lambda_1(x^2 + y^2 - 5)$
$1. \frac{\partial \mathcal{L}}{\partial x} = 2 - 2\lambda_1 x = 0$
$2. \frac{\partial \mathcal{L}}{\partial y} = -1 - 2\lambda_1 y = 0$
$3. \frac{\partial \mathcal{L}}{\partial \lambda_1} = x^2 + y^2 - 5 = 0$

Z 1. oraz 2. wiemy, że $\lambda_1 x = 1$ oraz $2\lambda_1 y = -1$, stąd $x = -2y$
Z 3. wiemy wtedy, że $5y^2 = 5 \Rightarrow y^2 = 1 \Rightarrow y = 1 \vee y = -1$

a) Przypadek, gdy $y = 1$
wtedy z 3. wiemy, że $x = 2 \vee x = -2$
gdy $y = 1$ i $x = 2$ wtedy z 2. widzimy że $\lambda_1 = -0.5$; podstawiając to do 1. mamy sprzeczność
gdy $y = 1$ i $x = -2$ wtedy z 2. widzimy że $\lambda_1 = -0.5$; punkt spełnia warunki

b) Przypadek, gdy $y = -1$
wtedy z 3. wiemy, że $x = 2 \vee x = -2$
gdy $y = -1$ i $x = 2$ wtedy z 2. widzimy że $\lambda_1 = 0.5$; punkt spełnia warunki
gdy $y = -1$ i $x = -2$ wtedy z 2. widzimy że $\lambda_1 = 0.5$; podstawiając to do 1. mamy sprzeczność

Mamy dwa punkty do analizy, $x_1 = [-2, 1]$, $\lambda_1 = -0.5$,$f(x_1) = -5$ oraz $x_2 = [2, -1]$,$\lambda_1 = 0.5$, $f(x_2) = 5$
Jakobian, czyli gradient ograniczenia to $[2x \;\; 2y]$
Przestrzeń zerowa dla x_1 to $J * Z = 0 \Rightarrow [-4 \; 2] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 0 \Rightarrow Z_{x_1} = \begin{bmatrix} z_1 \\ 2z_1 \end{bmatrix}$
Przestrzeń zerowa dla x_2 to $J * Z = 0 \Rightarrow [4 \; -2] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 0 \Rightarrow Z_{x_2} = \begin{bmatrix} z_1 \\ 2z_1 \end{bmatrix}$
Hesjan Lagranżjanu to $H_L = \begin{bmatrix} -2\lambda_1 & 0 \\ 0 & -2\lambda_1 \end{bmatrix}$, czyli dla x_1 to $H_L = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, a dla x_2 to $H_L = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
Określoność H_L dla x_1 to $[z_1 \; 2z_1] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ 2z_1 \end{bmatrix} = 5z_1^2$, więc punkt to minimum
Określoność H_L dla x_2 to $[z_1 \; 2z_1] \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ 2z_1 \end{bmatrix} = -5z_1^2$, więc punkt to maksimum

2.2 Optymalizacja – ograniczenia nierównościowe. Znajdź maksimum.
$f(x, y) = -x^2 + 2x + xy - y^2$ pod warunkiem
$g(x, y) = 2x + y \leq 2$

2.2

$\mathcal{L} = -x^2 + 2x + xy - y^2 - \lambda_1(2x + y - 2)$
$1. \frac{\partial \mathcal{L}}{\partial x} = -2x + 2 + y - 2\lambda_1 = 0$
$2. \frac{\partial \mathcal{L}}{\partial y} = x - 2y - \lambda_1 = 0$
$3. \lambda_1 \geq 0$ ($\lambda_1 = 0$ gdy $2x + y - 2 < 0$)

Przypadek a) ograniczenie 3. nie jest wiążące
Wtedy $\lambda_1 = 0$ i $2x + y - 2 < 0$
Z 2. widzimy, że $x = 2y$
Wtedy z 1. wiemy, że $-4y + 2 + y = 0 \Rightarrow 3y = 2 \Rightarrow y = \frac{2}{3}, x = \frac{4}{3}$
Widzimy że punkt nie spełnia warunku $2x + y - 2 < 0$

Przypadek b) ograniczenie 3. jest wiążące
Wtedy $\lambda_1 \geq 0$ i $2x + y = 2 \Rightarrow y = 2 - 2x$
Z 2. widzimy, że $x - 4 + 4x - \lambda_1 = 0 \Rightarrow 5x - 4 = \lambda_1$
Podstawiając lambdę oraz y do 1. mamy
$-2x + 2 + (2 - 2x) - 2(5x - 4) = 0 \Rightarrow -4x + 4 - 10x + 8 = 0 \Rightarrow x = \frac{6}{7}, y = \frac{2}{7}, \lambda_1 = \frac{2}{7}$
Punkt spełnia warunki.
Jakobian, tutaj gradient ograniczenia to $[2\ 1]$. Jego przestrzeń zerowa to
$J * Z = 0 \Rightarrow [2\ 1] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 0 \Rightarrow Z = \begin{bmatrix} z_1 \\ -2z_1 \end{bmatrix}$
Hesjan Lagranżjanu to $H_L = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$

Określoność powyższego H_L w przestrzeni zerowej Z to $[z_1 \ \ -2z_1] \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} z_1 \\ -2z_1 \end{bmatrix} = -14z_1^2$,

więc nasz punkt to maksimum

Praca domowa nr 3

Wybierz dowolną funkcję (najlepiej relatywnie łatwą w liczeniu), wybierz punkt początkowy oraz przyjmij inne założenia, jeżeli są potrzebne.
Przeprowadź:
1) Bracketing. Funkcja oraz wybrane punkty powinny pozwolić na przeprowadzenie minimum 1- 2 iteracji do znalezienia odpowiedniego przedziału zawierającego minimum.
2) Ternary search, minimum 2 iteracje, rozpoczynając od punktów uzyskanych z bracketingu.

Nie ma wymogu precyzji wyniku tego algorytmu, ważne aby poprawnie skrócił on przedział. W celu ułatwienia sobie zadania proszę pamiętać, że funkcja powinna mieć w analizowanym przedziale jedno minimum (powinna być funkcją unimodalną). Można korzystać z kalkulatora i innych pomocy, jeżeli chcemy wykonać zadanie na papierze. Można, ale nie trzeba, narysować wykresy przedstawiające kolejne kroki algorytmu. Można napisać program w R lub Pythonie który wykona to zadanie i opisze wykonywane kroki. Celem tego zadania jest lepsze zrozumienie działania bracketingu oraz ternary search.

In [5]:

```
install.packages("Deriv")
library(Deriv)

f<-function(x){(x)^2+10*x+2}
df<-Deriv(f,"x")
d2f<-Deriv(df,"x")
```

```
package 'Deriv' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Łukasz\AppData\Local\Temp\RtmpcfkFBq\downloaded_packages

Warning message:
"package 'Deriv' was built under R version 3.6.3"
```

In [6]:

```
band.search <- function(f, x0, delta) {
    points_band<-c()

    f.x0 <- f(x0)
    f.pd <- f(x0 + delta)
    f.md <- f(x0 - delta)
    points_band<-c(points_band,c(x0,x0 + delta,x0 - delta))
    if (!all(sapply(c(f.x0, f.pd, f.md), is.finite))) { #aby funkcja kontynuowała, wszy
stkie wartości muszą być skończone
        stop("Function value is not finite")
        }


    if (f.x0 <= min(f.pd, f.md)) { #Jeżeli w punkcie x0 wartość funkcji jest nie większ
a niż pozostałych dwóch punktach,
                                # to oznacza,  że w przedziale od
                                #x0 - delta do x0 + delta musi znajdować się poszukiwan
e minimum
        return(range(x0 + delta, x0 - delta))
    } else if (f.x0 > max(f.pd, f.md)) { # Jeżeli wartość funkcji w punkcie x0 jest wię
ksza niż w pozostałych dwóch punktach
                                        #oznacza to, że badana funkcja nie jest unimoda
lna w minimum
                                        #(ponieważ musi posiadać maksimum lokalne) i fu
nkcja przerywa działanie.
        stop("Function is not unimodal in minimum")
    } else if (f.pd > f.md) { #Ostatnim przypadkiem jest sytuacja, wktórej wartość funk
cji w punkcie x0 leży
                                #pomiędzy pozostałymi wartościami.
        search <- -delta #kierunek spadku funkcji
        f.x2 <- f.md

    } else {
        search <- delta #kierunek spadku funkcji
        f.x2 <- f.pd

    }


    x1 <- x0 #punkt startowy
    x2 <- x0 + search #punkt startowy + krok w kierunku spadku

    repeat { #powtarzamy aż do skutku

        x3 <- x0 + 2 * search # x3 to punkt startowy + 2 kroki w kierunku spadku
        points_band<-c(points_band,c(x1,x2,x3))

        if (!is.finite(x3)) {
            stop("Proper interval not found") #za każdym razem patrzymy czy wartość x n
owego jest skończona w rozumieniu przez R
        }

        f.x3 <- f(x3)

        if (!is.finite(f.x3)) {
            stop("Function value is not finite") #za każdym razem patrzymy czy wartość
 funkcji dla nowego kroku jest skończona
        }
```

```
        if (f.x3 > f.x2) { #idziemy tak daleko w kierunku spadku, aż wartość funkcji dl
a najdalszego krok przewyższy
                            #swoją wartością wartość funkcji dla drugiego kroku.
            result<-list("optimal interval"=range(x1, x3),"points band"=points_band)
            return(result)
        }

        search <- 2 * search #co pętlę podwajamy krok
        x1 <- x2 #co pętlę przesuwamy się z krokami dalej, w stronę spadku
        x2 <- x3
        f.x2 <- f.x3
    }
}
```

In [7]:

```
draw_iteration_band<-function(n,precision){
    x<-seq(min(result_band[[2]][3*n-2],result_band[[2]][3*n-1],result_band[[2]][3*n]),
           max(result_band[[2]][3*n-2],result_band[[2]][3*n-1],result_band[[2]][3*n]),p
recision)
    plot(x,f(x),lwd=0.01)
    points(result_band[[2]][3*n-2],f(result_band[[2]][3*n-2]),col="red",lwd=7)
    text(result_band[[2]][3*n-2],f(result_band[[2]][3*n-2])*1.007,"x1")
    points(result_band[[2]][3*n-1],f(result_band[[2]][3*n-1]),col="red",lwd=7)
    text(result_band[[2]][3*n-1],f(result_band[[2]][3*n-1])*1.007,"x2")
    points(result_band[[2]][3*n],f(result_band[[2]][3*n]),col="red",lwd=7)
    text(result_band[[2]][3*n],f(result_band[[2]][3*n])*1.005,"x3")
}
```

In [8]:

```
result_band=band.search(f,-2,0.5)
result_band
```

$`optimal interval`

-10 · -4

$`points band`

-2 · -1.5 · -2.5 · -2 · -2.5 · -3 · -2.5 · -3 · -4 · -3 · -4 · -6 · -4 · -6 · -10

In [9]:

```
ternary <- function(f, lower, upper, tol) {
    f.lower <- f(lower)
    f.upper <- f(upper)

    #kod do wykresu======================================================================
    brackets<-c()
    x_pairs<-c()
    brackets<-c(brackets,c(lower,upper))
    cat("Rozpoczęcie optymalizacji. Wartość dolna: ",lower,"Wartość górna: ",upper,"\n
\n")
    iteracja<-1
    #=======================================================================================
======

    while (abs(upper - lower) > 2 * tol) { #aż do momentu gdy różnica nie będzie mniejs
za niż 2xtolerancja
        cat("Rozpoczęcie pętli numer",iteracja,"Wartość dolna: ",lower,"Wartość górna:
 ",upper,"\n")
        x1 <- (2 * lower + upper) / 3 # zmienna x1 ląduje w 1/3 aktualnego przedziału
        f.x1 <- f(x1)
        x2 <- (lower + 2 * upper) / 3 # zmienna x2 ląduje w 2/3 aktualnego przedziału
        f.x2 <- f(x2)

        cat("Analizowane punkty w 1/3 i 2/3 przedziału: ",x1," oraz ",x2,"\n")
        cat("Wartość funkcji dla tych punktów to odpowiednio: ",f.x1," oraz ",f.x2,"\n"
)
        x_pairs<-c(x_pairs,c(x1,x2))

        if (f.x1 < f.x2) { #jeżeli wartość funkcji w lewej części jest niższa niż w pra
wej, ograniczamy przedział ze strony wyższej wartości
            upper <- x2
            f.upper <- f.x2
            cat("Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja warto
ści prawej strony przedziału na ",upper, "\n\n")
            brackets<-c(brackets,c(lower,upper))
        } else {            #jeżeli wartość funkcji w prawej części jest niższa niż w l
ewej,  ograniczamy przedział ze strony wyższej wartości
            lower <- x1
            f.lower <- f.x1
            cat("Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja warto
ści lewej strony przedziału na ",lower,"\n\n")
            brackets<-c(brackets,c(lower,upper))
        }
    iteracja<-iteracja+1
    }

    optim_point<-(upper + lower) / 2
    cat("Znaleziony punkt optymalny to: ",optim_point)
    result<-list("optim_point"=optim_point,"brackets"=brackets,"x_pairs"=x_pairs)
    return(result)
}
```

In [10]:

```
result=ternary(f,-10,-4,1e-3)
```

```
Rozpoczęcie optymalizacji. Wartość dolna:  -10 Wartość górna:  -4

Rozpoczęcie pętli numer 1 Wartość dolna:  -10 Wartość górna:  -4
Analizowane punkty w 1/3 i 2/3 przedziału:  -8  oraz  -6
Wartość funkcji dla tych punktów to odpowiednio:  -14  oraz  -22
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -8

Rozpoczęcie pętli numer 2 Wartość dolna:  -8 Wartość górna:  -4
Analizowane punkty w 1/3 i 2/3 przedziału:  -6.666667  oraz  -5.333333
Wartość funkcji dla tych punktów to odpowiednio:  -20.22222  oraz  -22.888
89
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -6.666667

Rozpoczęcie pętli numer 3 Wartość dolna:  -6.666667 Wartość górna:  -4
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.777778  oraz  -4.888889
Wartość funkcji dla tych punktów to odpowiednio:  -22.39506  oraz  -22.987
65
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.777778

Rozpoczęcie pętli numer 4 Wartość dolna:  -5.777778 Wartość górna:  -4
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.185185  oraz  -4.592593
Wartość funkcji dla tych punktów to odpowiednio:  -22.96571  oraz  -22.834
02
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.592593

Rozpoczęcie pętli numer 5 Wartość dolna:  -5.777778 Wartość górna:  -4.592
593
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.382716  oraz  -4.987654
Wartość funkcji dla tych punktów to odpowiednio:  -22.85353  oraz  -22.999
85
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.382716

Rozpoczęcie pętli numer 6 Wartość dolna:  -5.382716 Wartość górna:  -4.592
593
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.119342  oraz  -4.855967
Wartość funkcji dla tych punktów to odpowiednio:  -22.98576  oraz  -22.979
25
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.855967

Rozpoczęcie pętli numer 7 Wartość dolna:  -5.382716 Wartość górna:  -4.855
967
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.207133  oraz  -5.03155
Wartość funkcji dla tych punktów to odpowiednio:  -22.9571  oraz  -22.999
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.207133

Rozpoczęcie pętli numer 8 Wartość dolna:  -5.207133 Wartość górna:  -4.855
967
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.090078  oraz  -4.973022
Wartość funkcji dla tych punktów to odpowiednio:  -22.99189  oraz  -22.999
27
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.090078

Rozpoczęcie pętli numer 9 Wartość dolna:  -5.090078 Wartość górna:  -4.855
```

```
967
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.012041  oraz  -4.934004
Wartość funkcji dla tych punktów to odpowiednio:  -22.99986  oraz  -22.995
64
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.934004

Rozpoczęcie pętli numer 10 Wartość dolna:  -5.090078 Wartość górna:  -4.93
4004
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.038053  oraz  -4.986029
Wartość funkcji dla tych punktów to odpowiednio:  -22.99855  oraz  -22.999
8
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.038053

Rozpoczęcie pętli numer 11 Wartość dolna:  -5.038053 Wartość górna:  -4.93
4004
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.00337  oraz  -4.968687
Wartość funkcji dla tych punktów to odpowiednio:  -22.99999  oraz  -22.999
02
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.968687

Rozpoczęcie pętli numer 12 Wartość dolna:  -5.038053 Wartość górna:  -4.96
8687
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.014931  oraz  -4.991809
Wartość funkcji dla tych punktów to odpowiednio:  -22.99978  oraz  -22.999
93
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.014931

Rozpoczęcie pętli numer 13 Wartość dolna:  -5.014931 Wartość górna:  -4.96
8687
Analizowane punkty w 1/3 i 2/3 przedziału:  -4.999516  oraz  -4.984102
Wartość funkcji dla tych punktów to odpowiednio:  -23  oraz  -22.99975
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.984102

Rozpoczęcie pętli numer 14 Wartość dolna:  -5.014931 Wartość górna:  -4.98
4102
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.004655  oraz  -4.994378
Wartość funkcji dla tych punktów to odpowiednio:  -22.99998  oraz  -22.999
97
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.994378

Rozpoczęcie pętli numer 15 Wartość dolna:  -5.014931 Wartość górna:  -4.99
4378
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.00808  oraz  -5.001229
Wartość funkcji dla tych punktów to odpowiednio:  -22.99993  oraz  -23
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.00808

Rozpoczęcie pętli numer 16 Wartość dolna:  -5.00808 Wartość górna:  -4.994
378
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.003513  oraz  -4.998945
Wartość funkcji dla tych punktów to odpowiednio:  -22.99999  oraz  -23
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.003513

Rozpoczęcie pętli numer 17 Wartość dolna:  -5.003513 Wartość górna:  -4.99
```

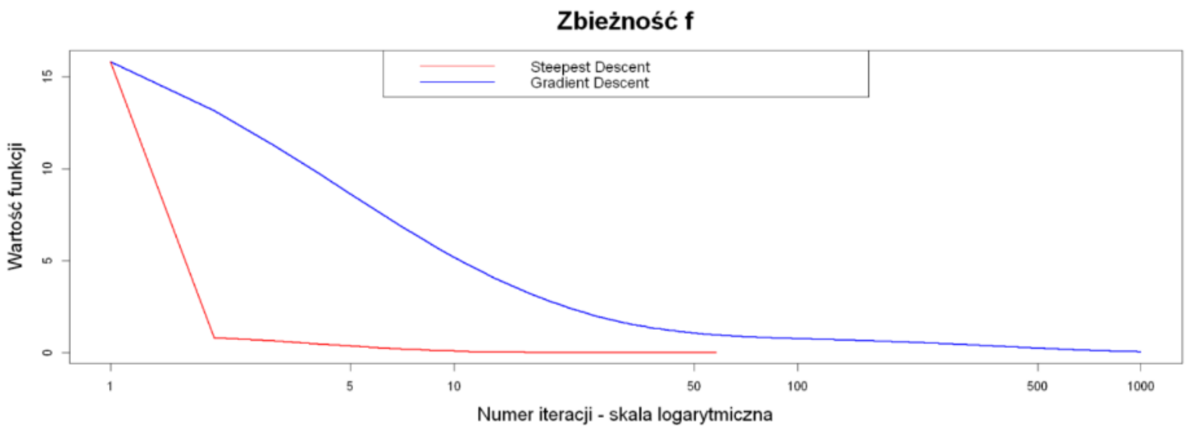
```
4378
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.000468  oraz  -4.997423
Wartość funkcji dla tych punktów to odpowiednio:  -23  oraz  -22.99999
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.997423

Rozpoczęcie pętli numer 18 Wartość dolna:  -5.003513 Wartość górna:  -4.99
7423
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.001483  oraz  -4.999453
Wartość funkcji dla tych punktów to odpowiednio:  -23  oraz  -23
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.001483

Rozpoczęcie pętli numer 19 Wartość dolna:  -5.001483 Wartość górna:  -4.99
7423
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.00013  oraz  -4.998776
Wartość funkcji dla tych punktów to odpowiednio:  -23  oraz  -23
Wartość funkcji z lewej strony niższa niż z prawej. Aktualizacja wartości
prawej strony przedziału na  -4.998776

Rozpoczęcie pętli numer 20 Wartość dolna:  -5.001483 Wartość górna:  -4.99
8776
Analizowane punkty w 1/3 i 2/3 przedziału:  -5.000581  oraz  -4.999679
Wartość funkcji dla tych punktów to odpowiednio:  -23  oraz  -23
Wartość funkcji z prawej strony niższa niż z lewej. Aktualizacja wartości
lewej strony przedziału na  -5.000581

Znaleziony punkt optymalny to:  -4.999679
```

Praca domowa nr 4

1. Opisz kluczowe Twoim zdaniem różnice pomiędzy metodami Gradient Descent a Steepest Descent.
2. Czym jest tzw. learning rate/alpha w ww. metodach, i jakie mogą być konsekwencje zbyt niskiej lub zbyt wysokiej wartości tego parametru?
3. Czym wytłumaczysz różnicę w zbieżności metod Gradient Descent i Steepest Descent?

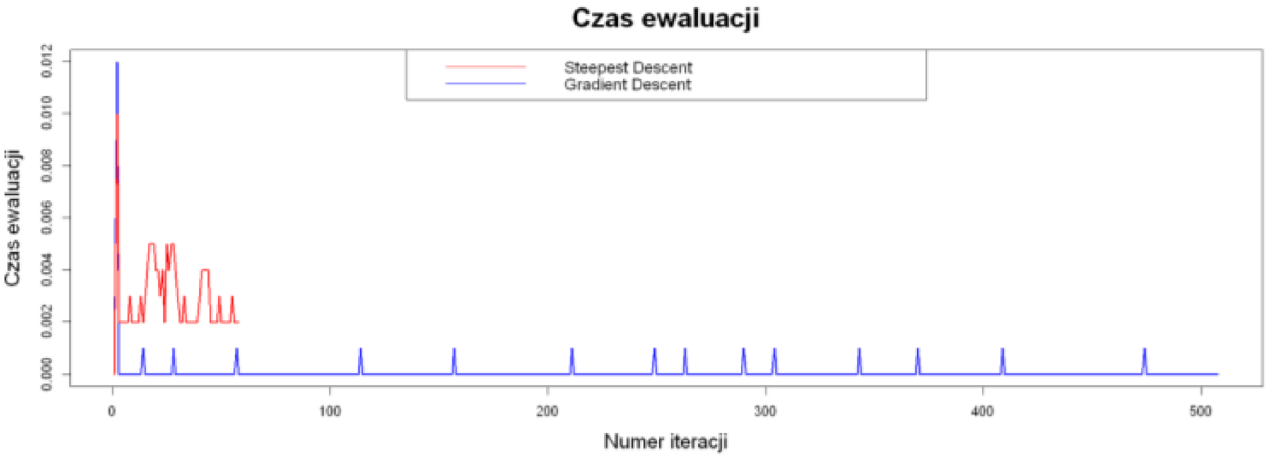
1. Gradient Descent omówiony na zajęciach za kierunek obiera wektor najszybszego spadku wartości funkcji, tzw. antygradient. Długością kroku w tej metodzie jest długość tego wektora pomnożona przez parametr alpha, czyli tzw. learning rate. W naszej implementacji Steepest Descent obiera analogiczny kierunek optymalizacji i początkową długość kroku (również opierając się na antygradiencie pomnożonym przez learning rate), jednak ta metoda <u>dodatkowo przeprowadza optymalizację wartości funkcji na odcinku</u> (łączącym początkowy punkt x0 z punktem x1 powstałym po odjęciu antygradientu pomnożonego przez learning rate). Ten optymalizacyjny podproblem może być rozwiązany w różny sposób, np. optymalizując funkcję jednej zmiennej korzystając z metody złotego podziału. Finalną długością kroku jest długość minimalizująca wartość funkcji na wspomnianym odcinku.
   Rozumieć należy, że <u>metoda Steepest Descent w ramach ww. podproblemu dokonuje optymalizacji parametru alpha</u> decydującego o długości kroku.

1. Learning rate jest hiperparametrem, arbitralnie obranym skalarem, przez który w omawianych przez nas metodach gradientowych mnoży się wektor gradientu (W metodzie Steepest Descent również dokonujemy doboru learning rate, jednak pamiętać należy, że właściwość ta wynika z konieczności implementacji tego rozwiązania). Źle dobrana wartość learning rate może spowodować problemy ze zbieżnością. Z jednej strony, niska wartość learning rate może być powodem długiego działania metody, lub przedwczesnego jej zakończenia, ponieważ długości kroków przy małych zmianach wartości funkcji będą niewielkie (niewielka długość wektora gradientu pomnożona przez niski learning rate), a ilość iteracji wysoka. Z drugiej strony, w przypadku metody Gradient Descent wysoka wartość tego parametru może powodować problemy ze zbieżnością, wielokrotne "oscylowanie" wokół wartości minimum.



Jak widać, metoda Steepest Descent cechuje się lepszą zbieżnością, ponieważ w każdej iteracji optymalizowana jest długość kroku w kierunku antygradientu, minimalizująca wartość funkcji w tym kierunku. Patrząc na wykres przez pryzmat implementacji metod na zajęciach, różnice w zbieżności pomiędzy metodami mogły również wyniknąć np. z doboru wartości learning rate dla poszczególnych metod, lub gdy właściwości analizowanej funkcji spowodowały, że optymalizacja długości kroku przyniosła dodatkową poprawę wartości funkcji.

1. Czym wytłumaczysz różnice w czasie wykonania?




```
Gradient Descent - czas ewaluacji:  0.02592325, optim x:  -1.517637e-09
, optim y:  0.0001560221
Steepest Descent - czas ewaluacji:  0.1685507, optim x:  -5.732683e-06
, optim y:  5.124782e-05
```

Steepest Descent w każdej iteracji musi dodatkowo rozwiązać podproblem optymalizacji wartości funkcji. Przekłada się to na dłuższy czas działania metody.

Praca domowa nr 5

1. Opisz własnymi słowami metodę funkcji kary.
2. (Dotyczy optymalizacji w metodzie funkcji kary na ćwiczeniach nr 5) Napisz jaką metodę optymalizacji funkcji jednej zmiennej wykorzystaliśmy na zajęciach do znalezienia optymalnej długości kroku dla metody najszybszego spadku.
3. Dlaczego kara zasadnicza za złamanie ograniczenia nierównościowego $g(x) \leq 0$ jest postaci $max(0, g(x))^2$, a kara za złamanie ograniczenia równościowego $h(x)^2$? Dlaczego wartość kary zasadniczej nie przyjmuje wartości odpowiednio $|max(0, g(x))|$ oraz $|h(x)|$?

1. Metoda funkcji kary polega na takim przekształceniu problemu optymalizacji funkcji z ograniczeniami, by rezultatem był ciąg problemów optymalizacji bez ograniczeń. Przekształcenie to polega na dodaniu dodatkowego komponentu stanowiącego funkcję kary, wpływającego na wartość funkcji w przedziałach w których ograniczenia nie są spełnione. W każdej kolejnej iteracji dotychczasowa wartość funkcji kary skalowana jest o ustalony mnożnik kary.

1. Wykorzystana została metoda złotego podziału

1. Funkcja kary powinna przyjmować wartość dodatnią w przypadku ograniczeń nierównościowych w postaci $g(x) \leq 0$ jedynie gdy ograniczenie zostaje złamane, czyli gdy wartość funkcji jest dodatnia. Gdy ograniczenie jest spełnione, chcemy by wartość funkcji kary wynosiła 0. Za złamanie ograniczenia równościowego chcemy nakładać karę, nieważne czy funkcja przyjmie wartości z jednej czy z drugiej strony ograniczenia równościowego. Wartość kary zasadniczej nie jest postaci $|max(0, g(x))|$ oraz $|h(x)|$, ponieważ taka postać kary nie jest różniczkowalna.

In [ ]:

```

```