Modele regresji i ich zastosowania
Labolatoria 11 i 12

Jan Solarz
243889

10 czerwca 2021

Spis treści

1 Regresogram i lokalne średnie

1.1 Zadanie 1

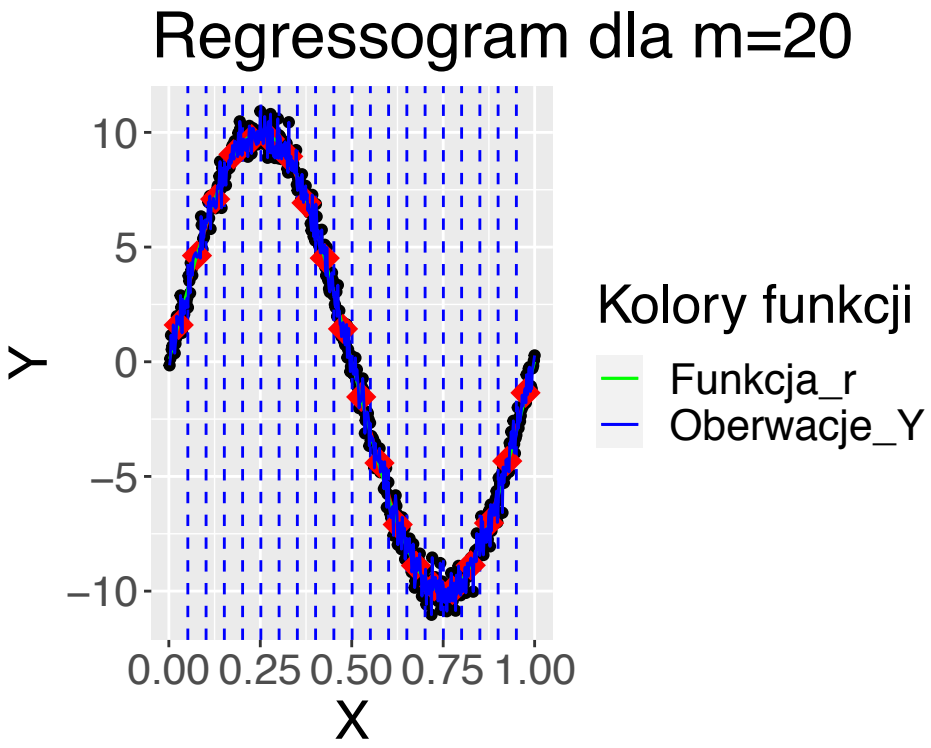
Generujemy 500 obserwacji gdzie $Y_i = r(xi) + \sigma\epsilon_i$,gdzie r to funkcja $r(x) = 10sin(2\pi x)$

```
library(ggplot2)
set.seed(123)
n=500
sigma=0.5
X <- c(1:n)/n
r<-10*sin(2*pi*X)
e<-rnorm(n,0,1)
Y<-r+sigma*e
```

1.2 Zadanie 2 i 3

Konstruujemy regresoramy dla dla różnych m przedziałów i rysujemy wykresy na jednym rysunku

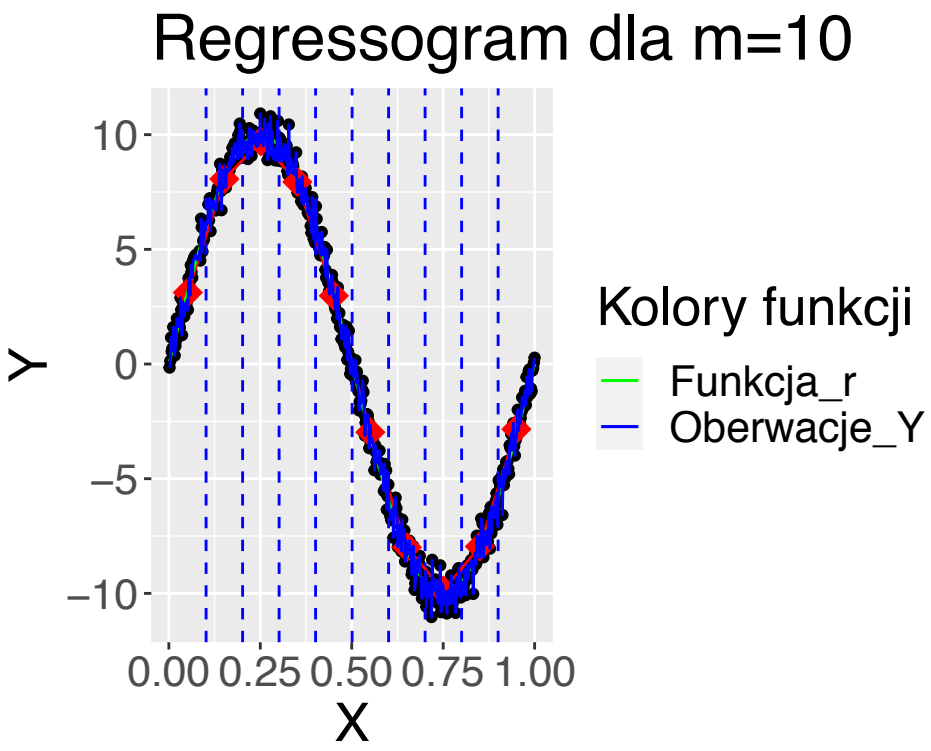
```
fig <- function(width, heigth){
     options(repr.plot.width = width, repr.plot.height = heigth)
}
fig(20,10)
library(HoRM)

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

regressogram(X, Y, nbins = 20, show.bins = TRUE,
             show.means = TRUE, show.lines = TRUE,
             x.lab = "X", y.lab = "Y", main = "Regressogram dla m=20") +geom_line(aes(y
  geom_line(aes(y = Y, colour = "Oberwacje_Y"))+
  scale_colour_manual(name="Kolory funkcji", values=c(Funkcja_r="green", Oberwacje_Y="bl
```



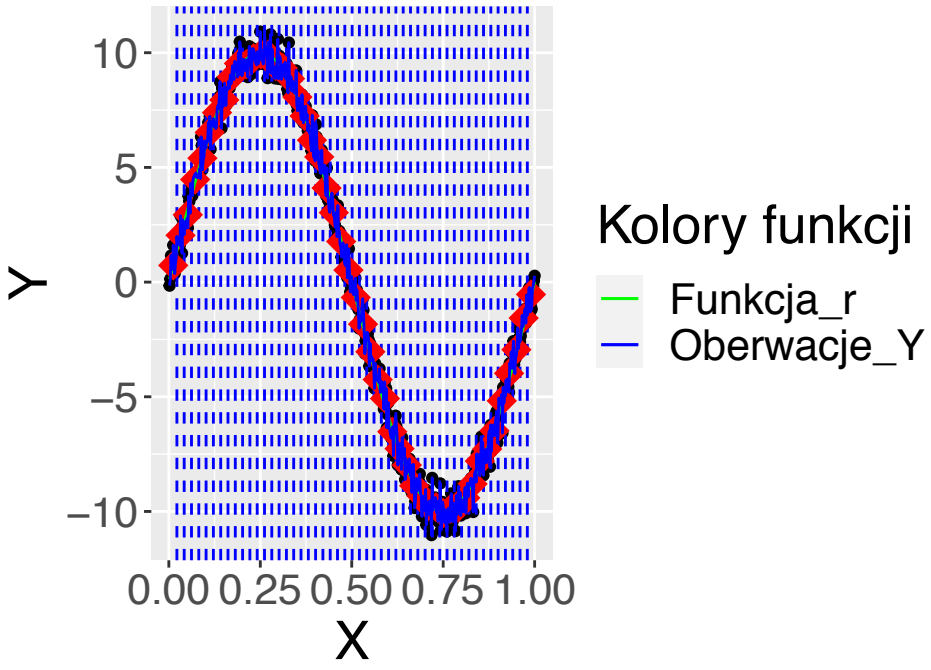
1.3 Zadanie 4

Wykonujemy te same symulacje dla m=10 i m=50

```
regressogram(X, Y, nbins = 10, show.bins = TRUE,
             show.means = TRUE, show.lines = TRUE,
             x.lab = "X", y.lab = "Y", main = "Regressogram dla m=10") +geom_line(aes(y
  geom_line(aes(y = Y, colour = "Oberwacje_Y"))+
  scale_colour_manual(name="Kolory funkcji", values=c(Funkcja_r="green", Oberwacje_Y="bl
```




```
regressogram(X, Y, nbins = 50, show.bins = TRUE,
             show.means = TRUE, show.lines = TRUE,
             x.lab = "X", y.lab = "Y", main = "Regressogram dla m=50") +geom_line(aes(y
  geom_line(aes(y = Y, colour = "Oberwacje_Y"))+
  scale_colour_manual(name="Kolory funkcji", values=c(Funkcja_r="green", Oberwacje_Y="bl
```

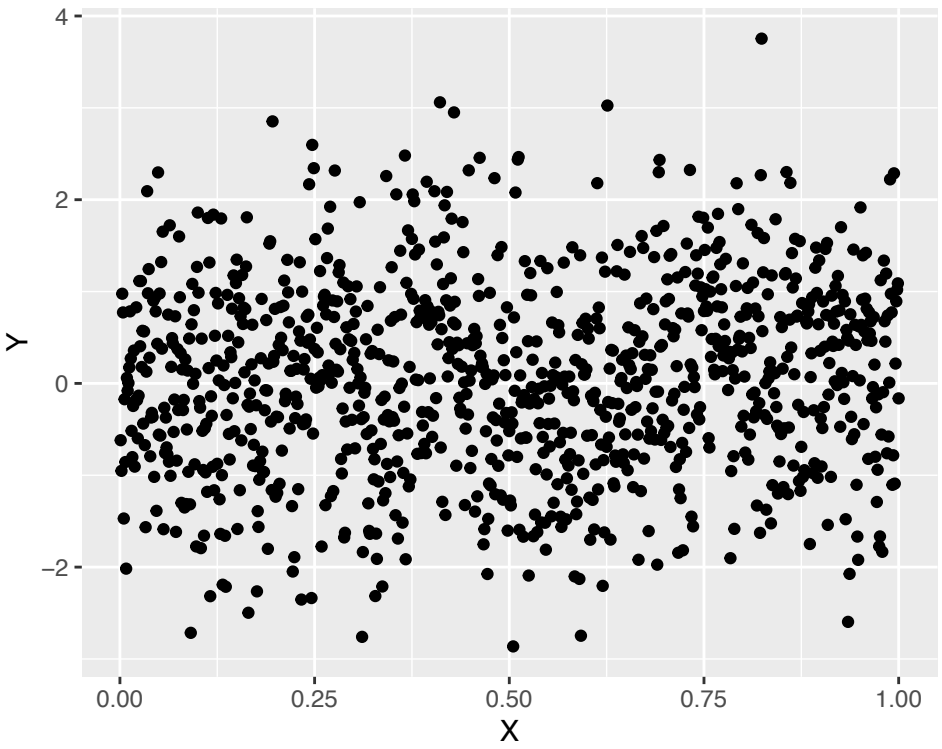


2 Estymator Nadaraya-Watsona i wielomiany lokalne

2.1 Zadanie 1

Generujemy 1000 obserwacji gdzie $Y_i = r(xi) + \sigma\epsilon_i$,gdzie r to funkcja Dopplera

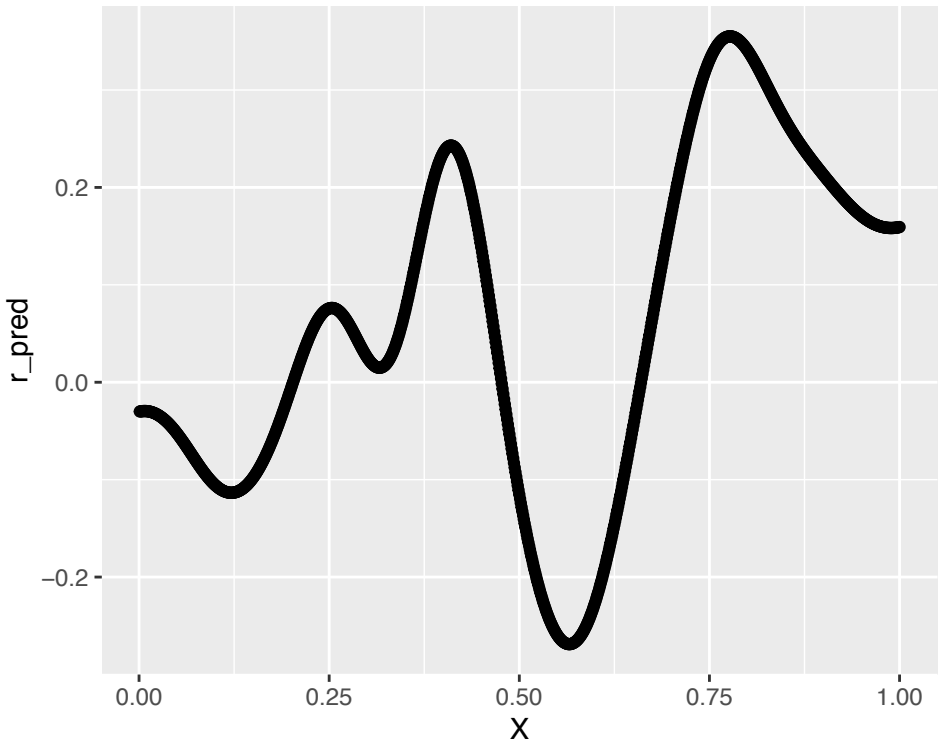
```
install.packages("CVThresh")

## Error in contrib.url(repos, "source"): trying to use CRAN without setting a mirror

library(CVThresh)

## Loading required package: wavethresh
## Loading required package: MASS
## WaveThresh: R wavelet software, release 4.6.8, installed
## Copyright Guy Nason and others 1993-2016
## Note: nlevels has been renamed to nlevelsWT
## Loading required package: EbayesThresh

n=1000
sigma=1
X <- c(1:n)/n
r<-dopp(X)
e<-rnorm(n,0,1)
r1<-sqrt(X*(1-X))*sin((2.1*pi)/(X+0.05))
Y<-r1+sigma*e


ggplot(data=data.frame(X,Y), aes(x=X,y=Y))+geom_point()
```



2.2 Zadanie 2

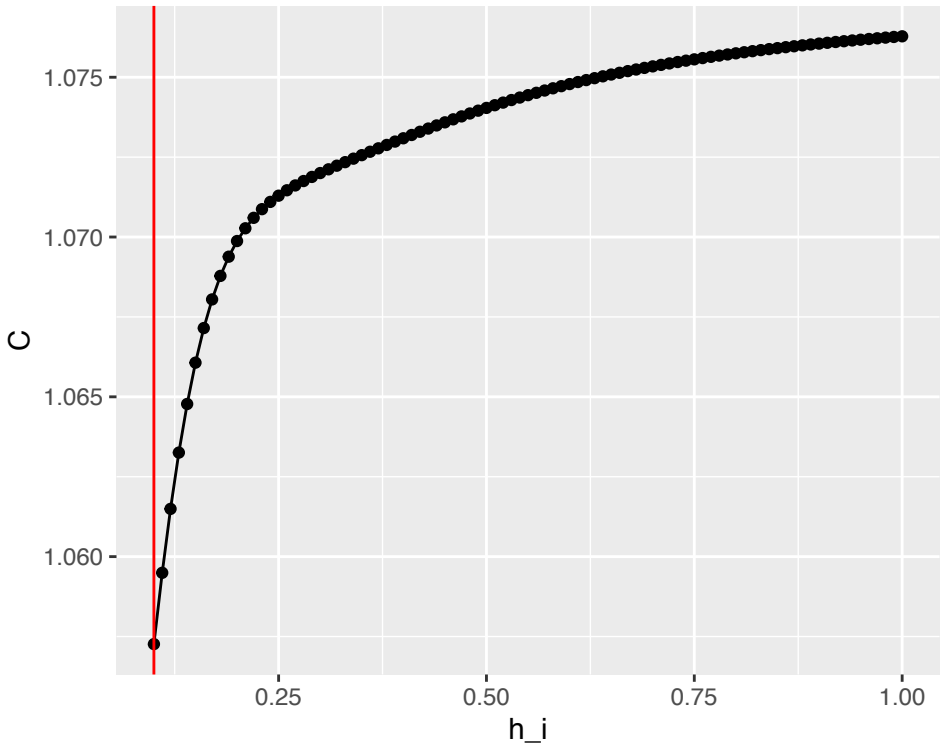
Konstruujemy estymator Nadaraya-Watsona oparty na jądrze gaussowskim. Estymujemy funkcje r, a następnie sprawdzamy optymalność parametru wagi h za pomocą metody LOOCF

```
#kernel
K<-function(x){
  dnorm(x)
}

#pred r function

#b<-NadarayaWatsonkernel(X, r1, h=0.05,gridpoint = X)
h=0.05
A<-outer(X,X,function(X,X1) K((X-X1)/h))
B<-A/rowSums(A)
r_pred<-B%*%Y
ggplot(data=data.frame(X,r_pred), aes(X,r_pred))+geom_point()
```




```
#loocv
X1<-c(1:n)/n
loocv<-function(h){
  A<-outer(X,X,function(X,X1) K((X-X1)/h))
  B<-A/rowSums(A)
  mean(((Y-B%*%Y)/(1-diag(B)))^2)
}
```

```
h_i<-seq(0.1,1,0.01)
C<-sapply(h_i,loocv)

h_best<-h_i[which.min(C)]
ggplot(data=data.frame(h_i,C), aes(x=h_i,y=C))+geom_line()+geom_point()+geom_vline(xinte
```

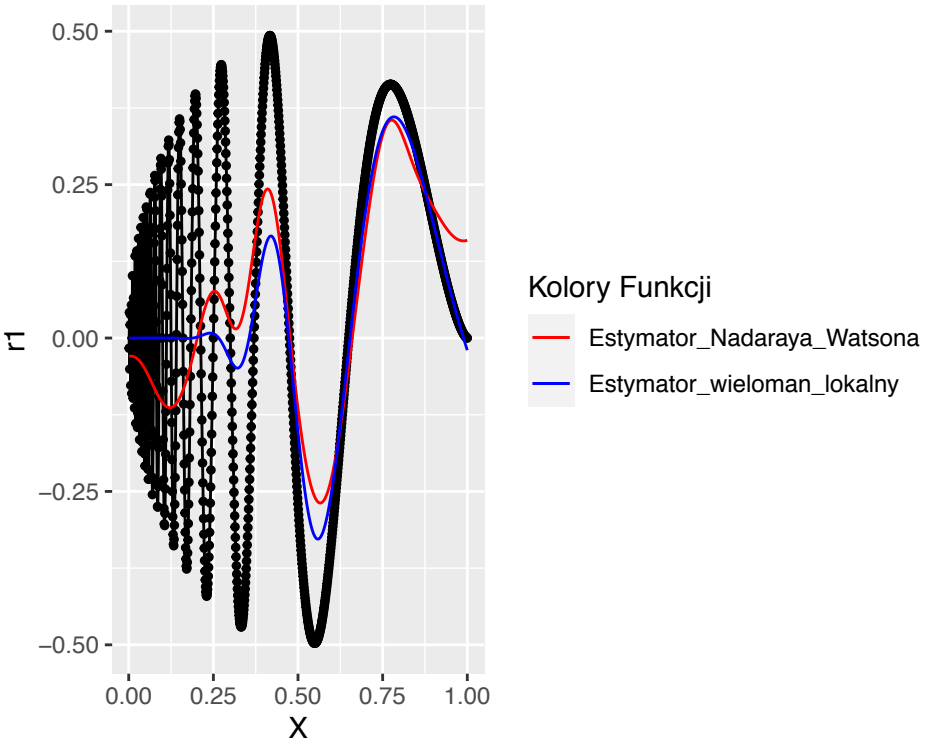


2.3 Zadanie 3 i 4

Konstruujemy estymator oparty na wielomianie lokalnym pierwszego stopnia. Następnie rysujemy oba estymatory i funkcje regreji Dopplera na jednym rysunku.

```
library(KernSmooth)

## KernSmooth 2.23 loaded
## Copyright M. P. Wand 1997-2009

l<-locpoly(X,r1,  degree =1, kernel = "normal",
        bandwidth=0.05,gridsize = 1000)




ggplot(data.frame(X,r1), aes(x = X, y = r1)) +
  geom_point(size = 1) +
  geom_line()+
  geom_line(aes(x=X, y = r_pred, colour="Estymator_Nadaraya_Watsona"))+
```

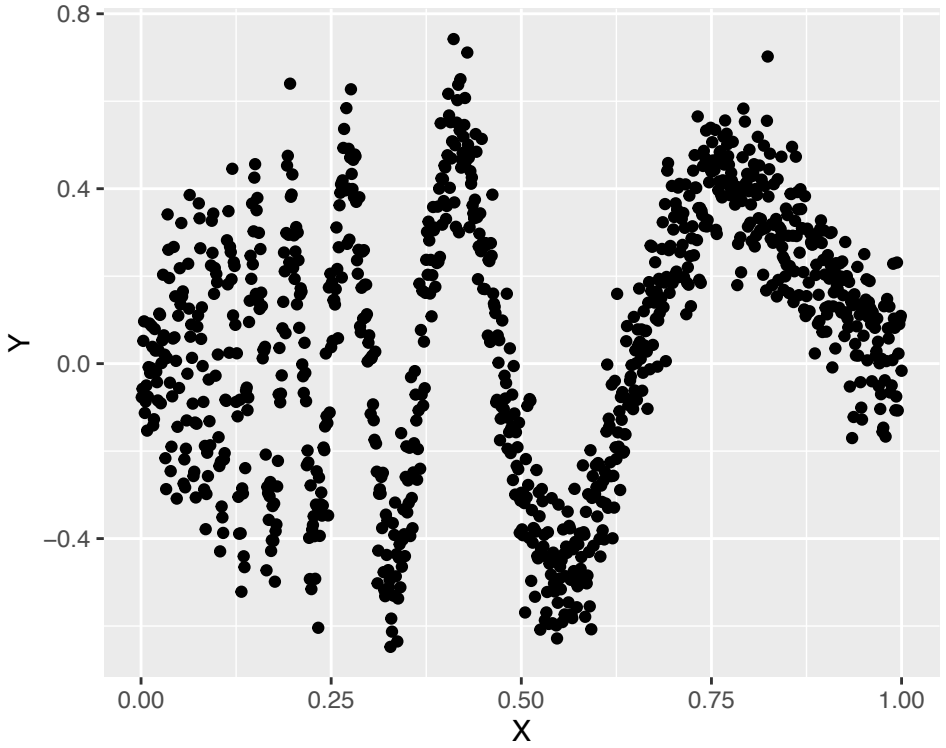
```
  geom_line(aes(x=X, y = l$y, colour="Estymator_wieloman_lokalny")) +
  scale_colour_manual(name="Kolory Funkcji",
      values=c(Estymator_Nadaraya_Watsona="red", Estymator_wieloman_lokalny="blue"))
```



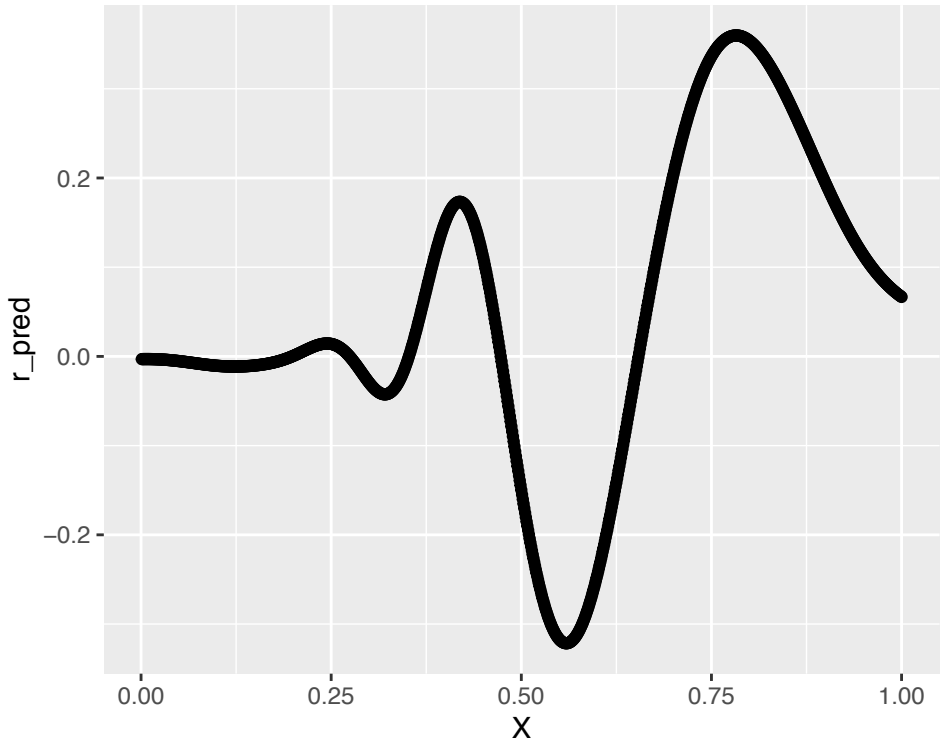
2.4 Zadanie 5

Powtórzenie symulacji dla $\sigma = 0.1$

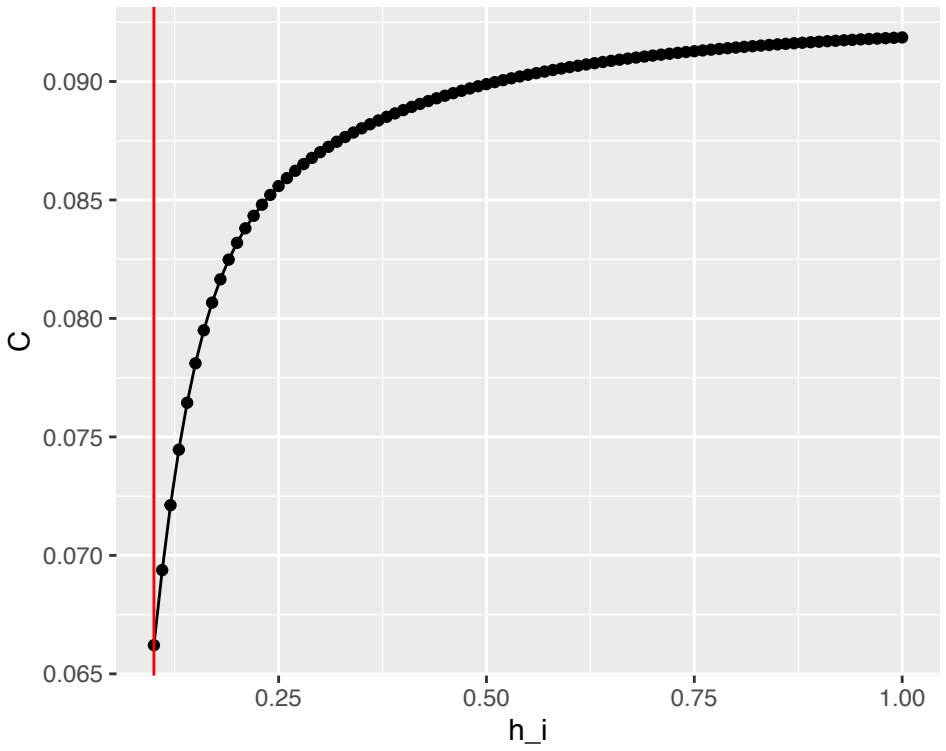
```
sigma=0.1
Y<-r1+sigma*e

library(KernSmooth)
l<-locpoly(X,r1,  degree =1, kernel = "normal",
        bandwidth=0.05,gridsize = 1000)
ggplot(data=data.frame(X,Y), aes(x=X,y=Y))+geom_point()
```



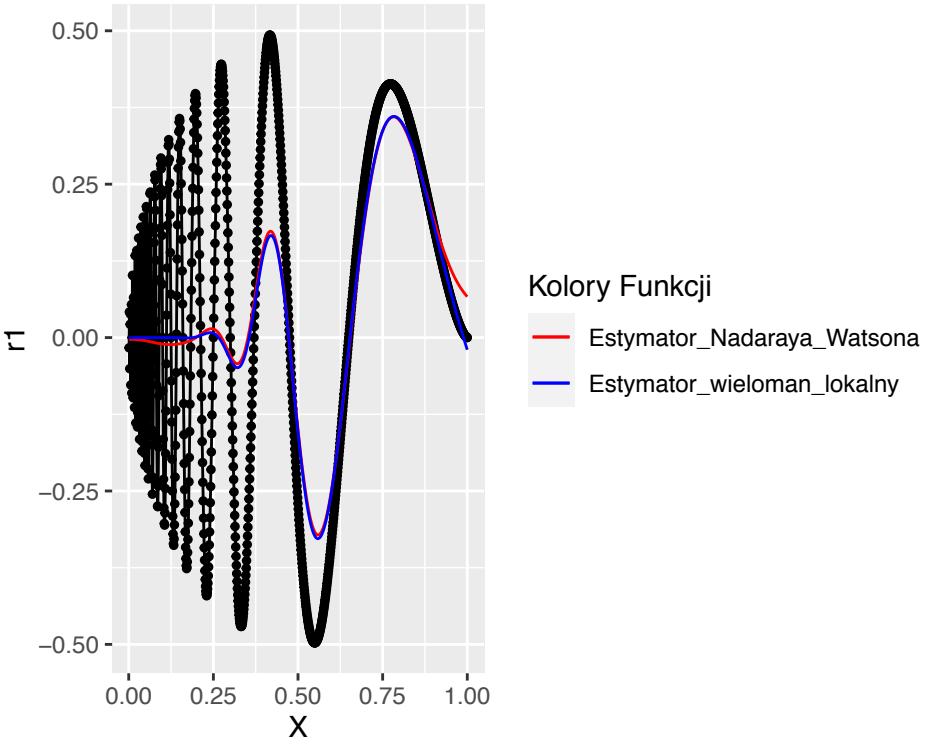

```
h=0.05
A<-outer(X,X,function(X,X1) K((X-X1)/h))
B<-A/rowSums(A)
r_pred<-B%*%Y
ggplot(data=data.frame(X,r_pred), aes(X,r_pred))+geom_point()
```



```r
#loocv
X1<-c(1:n)/n
loocv<-function(h){
  A<-outer(X,X,function(X,X1) K((X-X1)/h))
  B<-A/rowSums(A)
  mean(((Y-B%*%Y)/(1-diag(B)))^2)
}


h_i<-seq(0.1,1,0.01)
C<-sapply(h_i,loocv)

h_i[which.min(C)]

## [1] 0.1

ggplot(data=data.frame(h_i,C), aes(x=h_i,y=C))+geom_line()+geom_point()+geom_vline(xinte
```

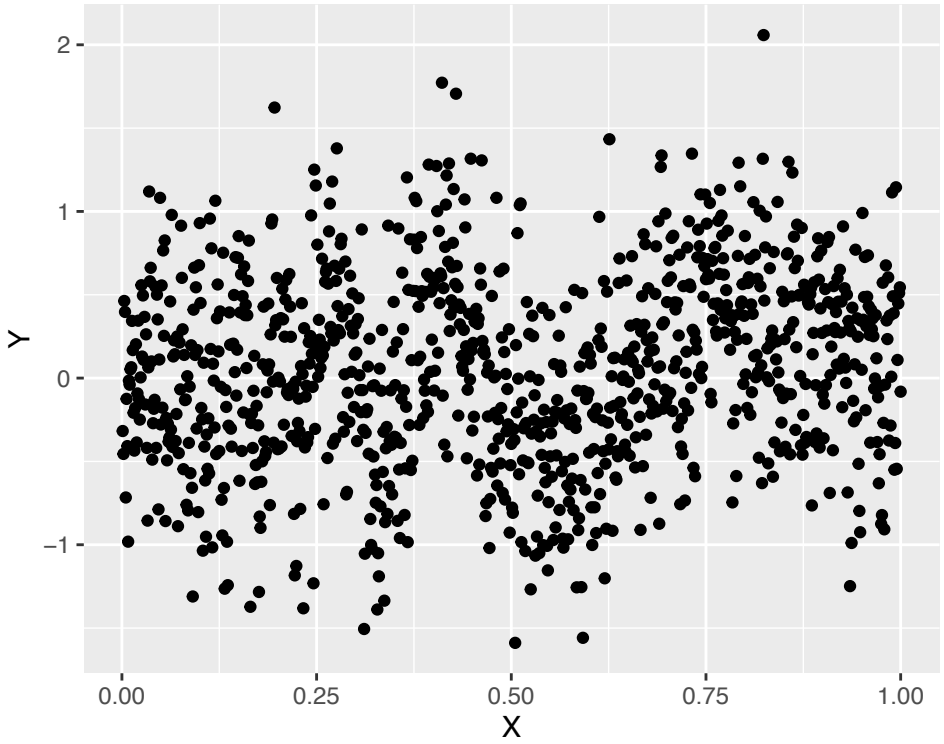



```r
ggplot(data.frame(X,r1), aes(x = X, y = r1)) +
  geom_point(size = 1) +
  geom_line()+
  geom_line(aes(x=X, y = r_pred, colour="Estymator_Nadaraya_Watsona"))+
  geom_line(aes(x=X, y = l$y, colour="Estymator_wieloman_lokalny")) +
  scale_colour_manual(name="Kolory Funkcji",
      values=c(Estymator_Nadaraya_Watsona="red", Estymator_wieloman_lokalny="blue"))
```

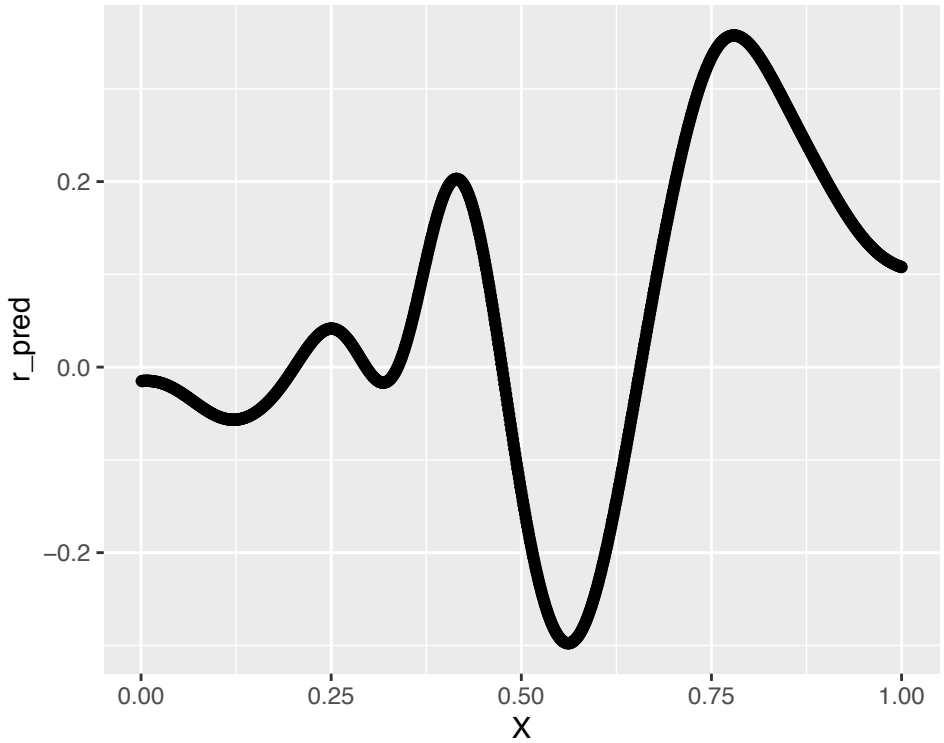


Powtórzenie symulacji dla $\sigma = 0.5$

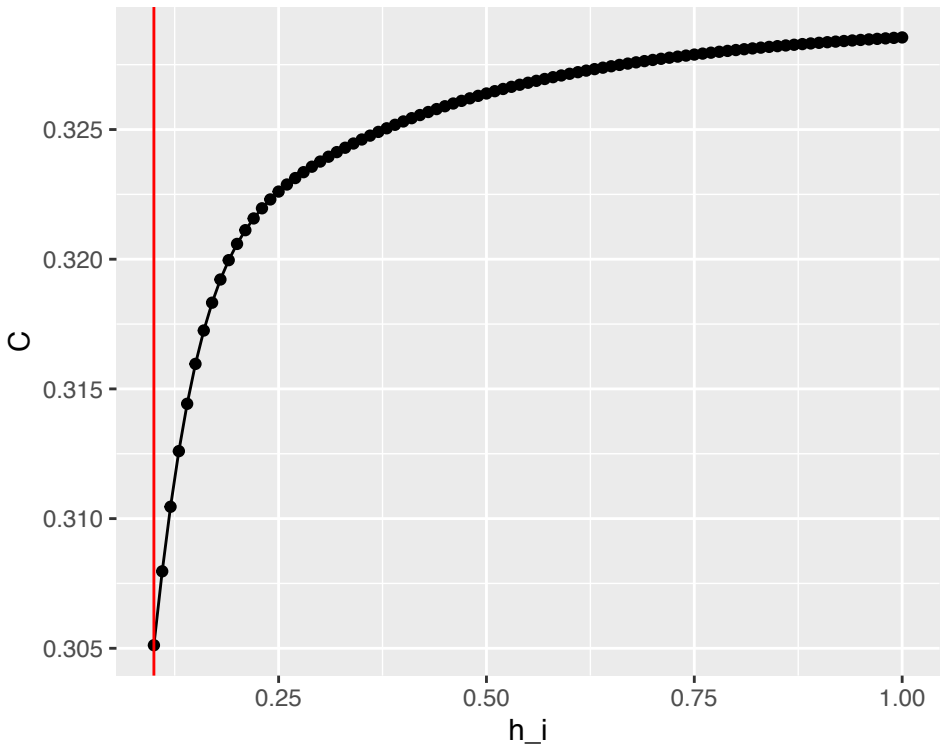
```
sigma=0.5
Y<-r1+sigma*e

library(KernSmooth)
l<-locpoly(X,r1,  degree =1, kernel = "normal",
        bandwidth=0.05,gridsize = 1000)
ggplot(data=data.frame(X,Y), aes(x=X,y=Y))+geom_point()
```



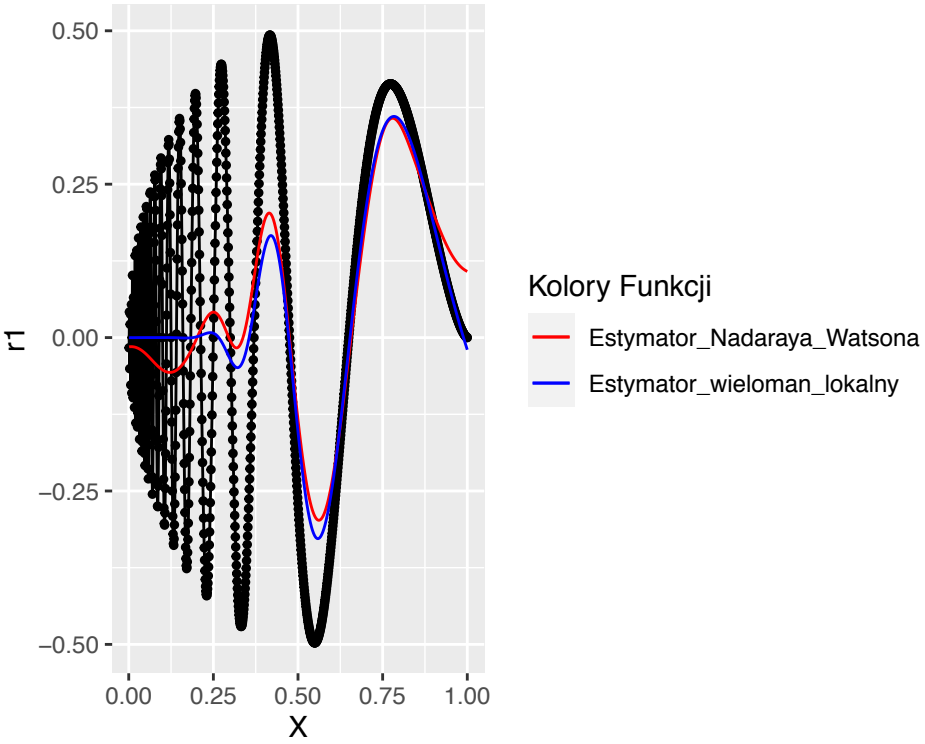

```
h=0.05
A<-outer(X,X,function(X,X1) K((X-X1)/h))
B<-A/rowSums(A)
r_pred<-B%*%Y
ggplot(data=data.frame(X,r_pred), aes(X,r_pred))+geom_point()
```



```r
#loocv
X1<-c(1:n)/n
loocv<-function(h){
  A<-outer(X,X,function(X,X1) K((X-X1)/h))
  B<-A/rowSums(A)
  mean(((Y-B%*%Y)/(1-diag(B)))^2)
}


h_i<-seq(0.1,1,0.01)
C<-sapply(h_i,loocv)

h_i[which.min(C)]

## [1] 0.1

ggplot(data=data.frame(h_i,C), aes(x=h_i,y=C))+geom_line()+geom_point()+geom_vline(xinte
```




```r
ggplot(data.frame(X,r1), aes(x = X, y = r1)) +
  geom_point(size = 1) +
  geom_line()+
  geom_line(aes(x=X, y = r_pred, colour="Estymator_Nadaraya_Watsona"))+
  geom_line(aes(x=X, y = l$y, colour="Estymator_wieloman_lokalny")) +
  scale_colour_manual(name="Kolory Funkcji",
      values=c(Estymator_Nadaraya_Watsona="red", Estymator_wieloman_lokalny="blue"))
```


0.50
0.25
0.00
−0.25
−0.50
r1
0.00
0.25
0.50
0.75
1.00
X
Kolory Funkcji
Estymator_Nadaraya_Watsona
Estymator_wieloman_lokalny