

# Podstawy Przetwarzania Danych

## Laboratoria 5: Redukcja zasumienia

dr inż. Marcin Luckner  
mluckner@mini.pw.edu.pl

Wydział Matematyki i Nauk Informatycznych

Wersja 1.1  
5 marca 2021

Projekt „NERW 2 PW. Nauka – Edukacja – Rozwój – Współpraca” współfinansowany jest ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Zadanie 10 pn. „Modyfikacja programów studiów na kierunkach prowadzonych przez Wydział Matematyki i Nauk Informatycznych”, realizowane w ramach projektu „NERW 2 PW. Nauka – Edukacja – Rozwój – Współpraca”, współfinansowanego jest ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

## Kryterium Chauveneta

```
1 Chauvenet <- function(datapoints){  
2   numdatapoints <- length(datapoints)  
3   dist <- abs(datapoints - mean(datapoints))/sd(  
4     datapoints)  
5   prob <- numdatapoints*dnorm(dist)  
6   sel <- prob >= 0.5  
7   datapoints <- datapoints[sel]  
8   return(datapoints)  
9 }
```

- Załączony skrypt, dostępny na stronie, stosuje kryterium Chauveneta w celu usunięcia obserwacji odstających.

# Zadanie 1

- Korzystając z załączonego skryptu sprawdź działanie kryterium Chauveneta na serii pomiarów.

3.8, 3.5, 3.9, 3.9, 3.4, 1.8

## Zadanie 2

- Ponów sprawdzanie kryterium Chauveneta na następujących seriach danych

3.8, 3.5, 3.9, 3.9, 3.4, 1.8, 1.8

3.8, 3.5, 3.9, 3.9, 3.4, 3.7, 3.5, 4.0, 3.3, 3.2, 100.0, 100.8

# Pakiet pracma

```
1 install.packages("pracma")  
2 library(pracma)
```

- Pakiet `pracma` zawiera funkcje dotyczące metod numerycznych, algebry liniowej, równań różniczkowych i analizy szeregów czasowych.

# Średnia krocząca

```
1 movavg(x, n, type=c("s", "t", "w", "m", "e", "r"))
```

- Metoda `movavg` wylicza średnią krocząca dla szeregu czasowego `x`.
- Parametr `n` określa szerokość okna.
- Parametr `type` określa algorytm wygładzania.

## Typy wygładzania

- s podstawowa średnia krocząca.
- t trójkątna średnia krocząca wylicza podstawową średnią krocząca z oknem  $\text{ceil}(n + 1)/2$ ; następnie wylicza drugą średnią krocząca na tych danych stosując ten sam rozmiar okna.
- w ważona średnia krocząca stosuje liniowe wagi do elementów z okna.
- m modyfikowana średnia krocząca. Pierwsza wartość jest wyliczana jak podstawowa średnia. Następne są wyliczane poprzez dodawanie nowej wartości i odjęcie poprzedniej średniej.
- e wykładnicza średnia krocząca waży elementy wykładniczo.
- r ruchoma średnia krocząca. Średnia wykładnicza zredukowana czynnikiem  $\frac{1}{n}$



## Zadanie 3

- Dla danych z pliku `distanceACC.csv` wykonaj wygładzenie danych podstawową średnią krocząca.
- Zastosuj rozmiary okna 3, 5 i 10.
- Wykreśl wygładzone dane.
- Zbadaj różnice w wyliczonych dla nich wartościach statystyk min, średnia i max.

## Zadanie 4

- Dla danych z pliku `distanceACC.csv` wykonaj wygładzenie danych z oknem 3.
- Zastosuj modyfikowaną, wykładniczą i trójkątną średnią krocząca.
- Wykreśl wygładzone dane.
- Zbadaj różnice w wyliczonych dla nich wartościach statystyk min, średnia i max.

## Tworzenie klasyfikatora One-SVM

```
1 library(e1071)
2 model <- svm(y~. , type='one-classification' , data=dat)
```

- Metoda `svm` buduje klasyfikator SVM.
- Parametr `type` pozwala ustalić typ klasyfikatora na `one-classification`.
- Parametr `data` określa zbiór danych.
- Formuła na początku wywołania definiuje zmienną okreśłą i zmienne określane.

## Określenie jądra klasyfikatora

```
1 model <- svm(y~. , type='one-classification' , data=dat ,  
  kernel="linear" )
```

- Parametr kernel pozwala określić jądro klasyfikatora.
- Dopuszczalne typy to:
  - linear,
  - polynomial,
  - radial,
  - sigmoid.

## Parametryzacja klasyfikatora

```
1 model <- svm(y~. , type='one-classification' , data=dat ,  
  kernel="radial" , nu=0.1, gamma=0.1)
```

- Parametr `nu` pozwala określić parametr klasyfikatora jednoklasowego.
- Jeżeli nie mamy do czynienia z jądrem liniowym to możemy też określić parametry jądra na przykład `gamma`

# Parametryzacja klasyfikatora

```
1 predict(model, dat)
```

- Funkcja `predict` zwraca wyniki klasyfikacji.
- Wyniki określają przypisanie obserwacji do klasy.

## Zadanie 5

- Dla danych z pliku `runningTime.csv` zbadaj działanie klasyfikatora jednoklasowego z jądrem liniowym i radialnym.
- Wykreśl uzyskane wyniki klasyfikacji

## Zadanie 6

- Dla danych z pliku `runningTime.csv` zbadaj działanie klasyfikatora jednoklasowego z jądrem radialnym z parametrami.
  - $\gamma \in \{0.1, 0.5\}$
  - $\nu \in \{0.1, 0.5\}$
- Wykreśl uzyskane wyniki klasyfikacji