

# Zagadnienie optymalnych kryteriów identyfikacji osoby na podstawie analizy mowy

Antoni Leon Dawidowicz<sup>1</sup>    Agata Trawińska<sup>2</sup>

<sup>1</sup>Uniwersytet Jagielloński  
w Krakowie

<sup>2</sup>Instytut Ekspertyz Sądowych  
w Krakowie

XXXV Konferencja  
Statystyka Matematyczna  
Wiśła 2009

Sformułowanie zagadnienia

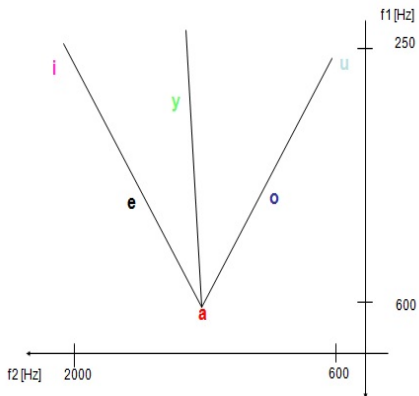
Propozycje rozwiązania

# Sformułowanie zagadnienia

Problem polega na identyfikacji osoby na podstawie jej wypowiedzi  
W każdej wypowiedzi można rozbić poszczególne samogłoski  
a,e,i,o,u na formanty

# Sformułowanie zagadnienia

Problem polega na identyfikacji osoby na podstawie jej wypowiedzi  
W każdej wypowiedzi można rozbić poszczególne samogłoski  
a,e,i,o,u na formanty



## Kiedy byłoby fajnie?

Gdyby każde wypowiedzenie danej samogłoski miało tę samą długość, wówczas można by było dla każdej samogłoski określić  $4n$ -wymiarowy wektor losowy.

## Gdzie są problemy?

- ▶ brakujące dane, co wynika głównie z jakości zapisu analizowanych wypowiedzi;
- ▶ różna liczebność wariantów poszczególnych samogłosek przy dostatecznie długiej wypowiedzi liczebność poszczególnych wariantów powinna odpowiadać frekwencyjności samogłosek właściwej dla języka polskiego;
- ▶ wybór treściowo tożsamy wypowiedzi trzech braci pozwolił natomiast zminimalizować problem zmienności układu czterech formantów, wynikającej z odmiennego prawo- i lewostronnego kontekstu fonetycznego samogłoski;

## Gdzie są problemy?

- ▶ brakujące dane, co wynika głównie z jakości zapisu analizowanych wypowiedzi;
- ▶ różna liczebność wariantów poszczególnych samogłosek przy dostatecznie długiej wypowiedzi liczebność poszczególnych wariantów powinna odpowiadać frekwencji samogłosek właściwej dla języka polskiego;
- ▶ wybór treściowo tożsamy wypowiedzi trzech braci pozwolił natomiast zminimalizować problem zmienności układu czterech formantów, wynikającej z odmiennego prawo- i lewostronnego kontekstu fonetycznego samogłoski;



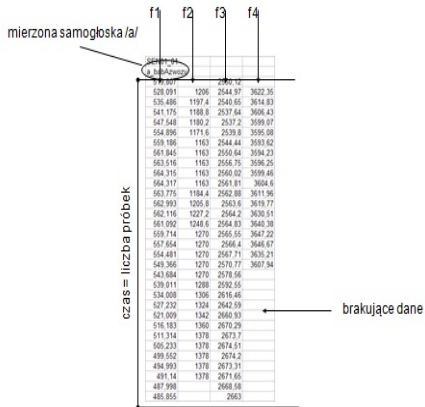
## Gdzie są problemy?

- ▶ brakujące dane, co wynika głównie z jakości zapisu analizowanych wypowiedzi;
- ▶ różna liczebność wariantów poszczególnych samogłosek przy dostatecznie długiej wypowiedzi liczebność poszczególnych wariantów powinna odpowiadać frekwencyjności samogłosek właściwej dla języka polskiego;
- ▶ wybór treściowo tożsamy wypowiedzi trzech braci pozwolił natomiast zminimalizować problem zmienności układu czterech formantów, wynikającej z odmiennego prawo- i lewostronnego kontekstu fonetycznego samogłoski;

## Gdzie są problemy?

- ▶ z kolei rejestracja ich wypowiedzi w sposób bezpośredni umożliwiła uwzględnienie czterech formantów z tą samą wagą istotności; rejestracja za pośrednictwem transmisji telefonicznej wymagałaby, z uwagi na charakterystykę pasma, przypisania formantom różnych wag relatywnie niższych dla formantów pierwszego i czwartego niż dla dwóch pozostałych, tj. formantów drugiego i trzeciego.

# Gdzie są problemy?



## Gdzie są problemy?

Głównym problemem, jaki pojawia się przy próbie opracowania algorytmy identyfikacji mówcy na podstawie wybranych formantów wybranych samogłosek jest fakt, że te same samogłoski wypowiedziane przez tego samego mówcę w tym samym kontekście mają różną długość

# Propozycje rozwiązania

Można ciąg pomiarów traktować, jako pomiary pewnej funkcji regresji. Jest to dość naturalne, gdyż każda wypowiedź samogłoski jest pewną funkcją czasu. Niektórzy autorzy sugereowali zastosowanie regresji wielomianowej.

# Propozycje rozwiązania

Można ciąg pomiarów traktować, jako pomiary pewnej funkcji regresji. Jest to dość naturale, gdyż każda wypowiedź samogłoski jest pewną funkcją czasu. Niektórzy autorzy sugereowali zastosowanie regrsji wielomianowej.

# Propozycje rozwiązania

Można ciąg pomiarów traktować, jako pomiary pewnej funkcji regresji. Jest to dość naturale, gdyż każda wypowiedź samogłoski jest pewną funkcją czasu. Niektórzy autorzy sugereowali zastosowanie regrsji wielomianowej.

# Propozycje rozwiązania

Regresja wielomianowa naszym zdaniem nie jest adekwatna do problemu Bardziej naturalna jest procedura nieparametryczna

$$\hat{f}(x) = \frac{\sum_{k=1}^n y_k K\left(\frac{x-x_k}{h}\right)}{\sum_{k=1}^n K\left(\frac{x-x_k}{h}\right)}$$

gdzie  $K$  jest pewną krzywą dzwonową



# Propozycje rozwiązania

Regresja wielomianowa naszym zdaniem nie jest adekwatna do problemu Bardziej naturalna jest procedura nieparametryczna

## Estymator Nadarayi-Watsona

$$\hat{f}(x) = \frac{\sum_{k=1}^n y_k K\left(\frac{x-x_k}{h}\right)}{\sum_{k=1}^n K\left(\frac{x-x_k}{h}\right)}$$

gdzie  $K$  jest pewną krzywą dzwonową

# Propozycje rozwiązania

Regresja wielomianowa naszym zdaniem nie jest adekwatna do problemu Bardziej naturalna jest procedura nieparametryczna

## Estymator Nadarayi-Watsona

$$\hat{f}(x) = \frac{\sum_{k=1}^n y_k K\left(\frac{x-x_k}{h}\right)}{\sum_{k=1}^n K\left(\frac{x-x_k}{h}\right)}$$

gdzie  $K$  jest pewną krzywą dzwonową

# Propozycje rozwiązania

Regresja wielomianowa naszym zdaniem nie jest adekwatna do problemu Bardziej naturalna jest procedura nieparametryczna

## Estymator Nadarayi-Watsona

$$\hat{f}(x) = \frac{\sum_{k=1}^n y_k K\left(\frac{x-x_k}{h}\right)}{\sum_{k=1}^n K\left(\frac{x-x_k}{h}\right)}$$

gdzie  $K$  jest pewną krzywą dzwonową

## Propozycje rozwiązania

Przed zastosowaniem estymatora Nadarayi-Watsona można dokonać takiej zamiany zmiennych tak, by jednostką czasu był okres wypowiedzenia samogłoski, czyli uwolnić się od problemów związanych z różnym czasem wypowiedzenia tej samej samogłoski. Rzecz jasna należy odpowiednio zrelatywizować szerokość pasma  $h$ .

## Propozycje rozwiązania

Przed zastosowaniem estymatora Nadarayi-Watsona można dokonać takiej zamiany zmiennych tak, by jednostką czasu był okres wypowiedzenia samogłoski, czyli uwolnić się od problemów związanych z różnym czasem wypowiedzenia tej samej samogłoski. Rzecz jasna należy odpowiednio zrelatywizować szerokość pasma  $h$ .

Dziękujemy za uwagę