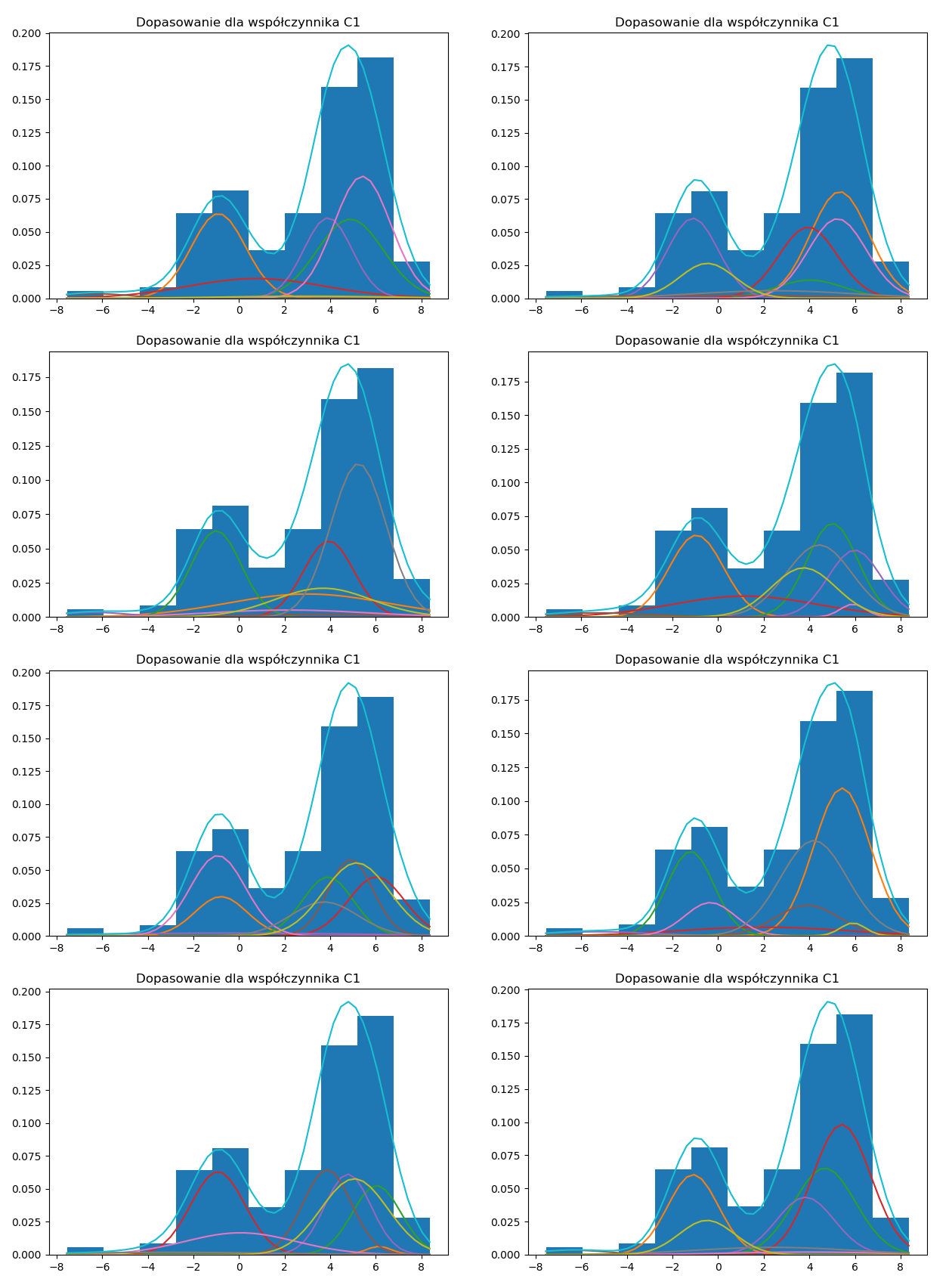
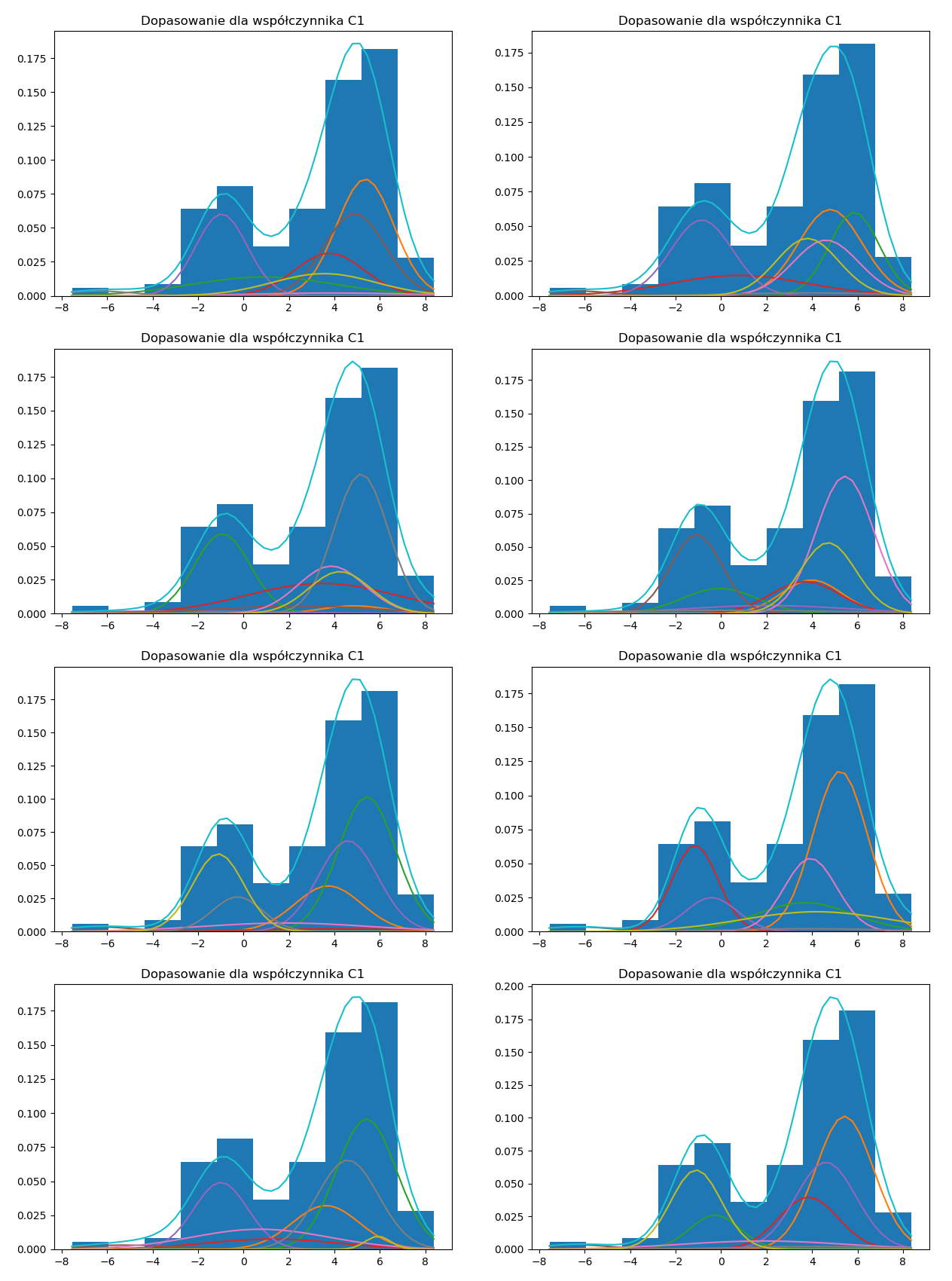
# Zadanie 1 – Model GMM

## Wykresy dla różnej liczby iteracji algorytmu EM

Rysunek 1. Dopasowania dla różnej ilości iteracji algorytmu EM (1-8)



Rysunek 2.Dopasowania dla różnej wartości iteracji algorytmu EM (10-80 w kroku co 10)

Analizując wykresy dla różnej wartości iteracji algorytmu EM można dostrzec iż dla niższych wartości dopasowania nie zawsze pasują do histogramu. Jest to kwestia w pełni losowa czy dopasowane krzywe będą dobrze odzwierciedlać daną cechę, czy też będą nie dość dobre(tak jak widać wyżej na wykresie dla 2 i 7 współczynników). Im większa liczba iteracji tym wykresy robią się coraz bardziej podobne do siebie, oraz do wartości dla konkretnego współczynnika.

## Analiza wartości kryterium informacyjnego w zależności od liczby iteracji

Tabela . Wartości kryterium informacyjnego w zależności od liczby iteracji algorytmu EM

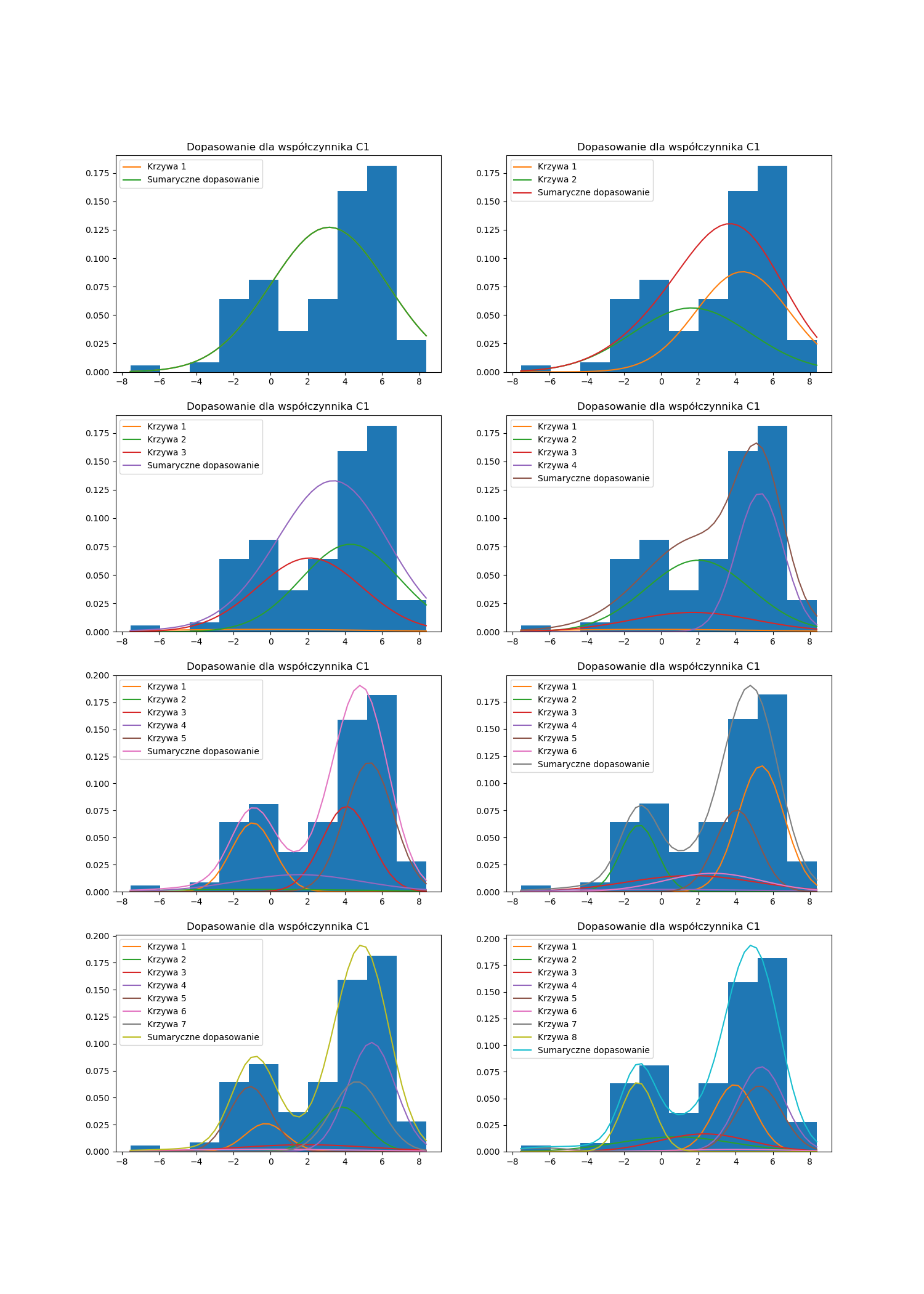
|  |  |  |
| --- | --- | --- |
| Liczba iteracji | AIC | BIC |
| 1 | 13599 | 14334 |
| 3 | 13576 | 14311 |
| 5 | 13449 | 14183 |
| 7 | 13476 | 14210 |
| 9 | 13421 | 14155 |
| 10 | 13528 | 14262 |
| 20 | 13404 | 14138 |
| 30 | 13491 | 14226 |
| 50 | 13497 | 14232 |
| 70 | 13445 | 14179 |
| 90 | 13492 | 14227 |

Stwierdzam, iż jako kryterium przyjmuję AIC, gdyż różnice w wartości parametru AIC większe niż 10 gwarantują iż dopasowanie modelu będzie dobre. Poza tym różnice w wartości parametru BIC są zdecydowanie mniejsze niż dla AIC. Dodatkowo BIC jest parametrem, który z definicji radzi sobie lepiej w opisywaniu modeli prostych.

## Ile komponentóW jest optymalnych dla modelowania głosek

Rysunek 3. Dopasowania dla różnej liczby komponentów (1-8)

asd

Zgodnie z powyższymi wykresami, można dostrzec, iż liczba komponentów wystarczających do poprawnego zamodelowania głosek to 6. Każda mniejsza wartość sprawa iż sumaryczne dopasowanie nie jest na tyle poprawnym aby dobrze zamodelować współczynniki(w powyższym przypadku współczynnik C1).

# Zadanie 2 – Log-Prawdopodobieństwa

## Obliczone wartości log-prawdopodobieństw

Tabela 2. Wartości log-prawdopodobieństw

|  |  |
| --- | --- |
|  | -28,99 |
|  | -257,22 |
|  | -29,93 |
|  | -336,76 |

Uzyskane wartości są zgodne z oczekiwaniami. Pokazują iż nagranie głoski /a/ pasuje do modelu dla tej głoski, a nagranie głoski /i/ pasuje do modelu głoski /i/. Wartości prawdopodobieństw odwrotnych (jak nagranie głoski /a/ pasuje do modelu głoski /i/ oraz na odwrót) są dużo niższe, co wskazuje na to że dopasowanie głoski /a/ do modelu /i/ oraz na odwrót jest bardzo mało prawdopodobne.

## Wartości prawdopodobieństw z reguły bayesa dla nowo utworzonych nagrań

Nowe nagrania uzyskują bardzo wysokie prawdopodobieństwo zaklasyfikowania jako odpowiednia głoska. Dla próby przypasowania do niepasującego modelu prawdopodobieństwo jest tak marginalnie małe, że można uznać iż wynosi zero.