

# Laboratorium 4

(4+2)

## 1. Wybór modelu

W pliku regr.dat znajdują się wektory  $x$  i  $y$ . Proszę dopasować kolejno wielomiany stopnia 2-7 do zależności  $(x_i, y_i)$ . Wiemy, że niepewność zmiennej  $y$  dla każdego punktu pomiarowego wynosi 2.

- Obliczyć  $\chi^2$  na stopień swobody dla każdego wielomianu.
- Obliczyć funkcję kosztu dla kryterium AIC (Akaike Information Criterion)

$$-2\ln(\text{funkcja największej wiarygodności}) + 2(\text{liczba parametrów})$$

w naszym przypadku możemy wziąć

$$\chi^2 + 2(\text{liczba parametrów})$$

Proszę wyliczyć względne prawdopodobieństwo modeli względem modelu o najmniejszym koszcie AIC

$$\frac{P_i}{P_j} = \exp(-(AIC_i - AIC_j)/2)$$

Proszę powtórzyć to samo dla BIC (Bayesian Information Criterion)

$$\chi^2 + (\text{liczba parametrów})\ln(\text{liczba punktów})$$

- Oszacować jakość dopasowania metodą “krosvalidacji” (metoda krzyżowa). Pary punktów  $(x_i, y_i)$  podzielić na dwie części, zbiór trenujący (wszystkie punkty bez jednego) i zbiór testujący (pozostały, jeden punkt). Wykonać dopasowanie wielomianów  $f(x)$  stopnia 2-7 na zbiorze trenującym. Na zbiorze testującym (punkcie) obliczyć odpowiednie odchylenia kwadratowe

$$(y_k - f(x_k))^2$$

Proszę te obliczenia wykonać zmieniając zbiór (punkt) testujący i uśrednić odchylenie dla wszystkich zbiorów testujących. Porównać oszacowane odchylenie dopasowania w zależności od stopnia wielomianu.