Wydział Elektrotechniki, Informatyki i Telekomunikacji
Uniwersytet Zielonogórski
Instytut Sterowania i Systemów Informatycznych

**Laboratorium metod i technik optymalizacji**

# Regresja nieliniowa

Rozważmy sytuację, gdy obserwacje pewnego obiektu można przedstawić w postaci

$$y_i = f(x_i; a_0, a_1, \ldots, a_m) + e_i, \quad i = 1, \ldots, n,$$

gdzie: $y_i$ – zaobserwowana wartość zmiennej objaśnianej, $f(x_i; a_0, a_1, \ldots, a_m)$ – funkcja zmiennej objaśniającej $x_i$ i nieznanych stałych parametrów $a_0, a_1, \ldots, a_m$, oraz $e_i$ – błąd przypadkowy. Przykładową charakterystyką statyczną dla $m = 1$ może być

$$f(x; a_0, a_1) = a_0(1 - e^{-a_1 x}).$$

Dla uproszczenia, załóżmy dalej właśnie $m = 1$ oraz przyjmijmy notację

$$f(x_i) = f(x_i; a_0, a_1),$$

co oznacza, że obserwacje można teraz zapisać w postaci

$$y_i = f(x_i) + e_i, \quad i = 1, \ldots, n.$$

Problem polega na znalezieniu ocen parametrów $a_0$ i $a_1$ minimalizujących sumę kwadratów residuów

$$S_r(a_0, a_1) = \sum_{i=1}^{n} (y_i - f(x_i; a_0, a_1))^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2.$$

Można go rozwiązać w oparciu o klasyczną metodę najmniejszych kwadratów dla regresji liniowej. Pomysł jest następujący: Załóżmy, że $a_{0,j}$ oraz $a_{1,j}$ są przybliżeniami rzeczywistych wartości $a_0$ oraz $a_1$ na początku $j$-tego kroku. Korzystając z rozwinięcia w szereg Taylora, dokonujemy linearyzacji funkcji $f$ w otoczeniu punktu $(a_{0,j}, a_{1,j})$ i pomijamy człony wyższych rzędów:

$$f(x_i)_{j+1} = f(x_i)_j + \frac{\partial f(x_i)_j}{\partial a_0} \Delta a_0 + \frac{\partial f(x_i)_j}{\partial a_1} \Delta a_1,$$

gdzie:

$$\begin{aligned}
\frac{\partial f(x_i)_j}{\partial a_0} &= \frac{\partial f(x_i; a_{0,j}, a_{1,j})}{\partial a_0}, \quad \Delta a_0 = a_{0,j+1} - a_{0,j}, \\
\frac{\partial f(x_i)_j}{\partial a_1} &= \frac{\partial f(x_i; a_{0,j}, a_{1,j})}{\partial a_1}, \quad \Delta a_1 = a_{1,j+1} - a_{1,j}.
\end{aligned}$$

W rezultacie otrzymujemy model liniowy ze względu na nieznane parametry (są nimi teraz $\Delta a_0$ oraz $\Delta a_1$):

$$y_i - f(x_i)_j = \frac{\partial f(x_i)_j}{\partial a_0}\Delta a_0 + \frac{\partial f(x_i)_j}{\partial a_1}\Delta a_1 + e_i, \quad i = 1, \ldots, n.$$

W notacji macierzowej, powyższe równania przyjmują postać

$$\{D_j\} = \{Z_j\}\{\Delta A\} + \{E_j\},$$

gdzie

$$\{Z_j\} = \begin{bmatrix} \partial f(x_1)_j/\partial a_0 & \partial f(x_1)_j/\partial a_1 \\ \partial f(x_2)_j/\partial a_0 & \partial f(x_2)_j/\partial a_1 \\ \vdots & \vdots \\ \partial f(x_n)_j/\partial a_0 & \partial f(x_n)_j/\partial a_1 \end{bmatrix}, \quad \{D_j\} = \begin{bmatrix} y_1 - f(x_1)_j \\ y_2 - f(x_2)_j \\ \vdots \\ y_n - f(x_n)_j \end{bmatrix}$$

$$\{\Delta A\} = \begin{bmatrix} \Delta a_0 \\ \Delta a_1 \end{bmatrix}, \qquad \{E_j\} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Zastosowanie do powyższego układu równań liniowych poznanej już metody najmniejszych kwadratów prowadzi do *układu równań normalnych*

$$\left[\{Z_j\}^T\{Z_j\}\right]\{\Delta A\} = \{Z_j\}^T\{D_j\},$$

który rozwiązuje się ze względu na wektor $\{\Delta A\}$. Na tej podstawie otrzymuje się nowe oceny

$$a_{0,j+1} = a_{0,j} + \Delta a_0, \quad a_{1,j+1} = a_{1,j} + \Delta a_1,$$

dla których opisane obliczenia powtarza się. Proces przerywa się gdy kryterium stopu

$$|\varepsilon_a| = \min\left\{\left|\frac{a_{0,j+1} - a_{0,j}}{a_{0,j+1}}\right|, \left|\frac{a_{1,j+1} - a_{1,j}}{a_{1,j+1}}\right|\right\}$$

stanie się mniejsze od zadanej dokładności. Opisana procedura nosi nazwę *metody Gaussa-Newtona*.

Ćwiczenie polega na zastosowaniu metody Gaussa-Newtona do rozwiązania następujących zadań:

1. Do danych

| $x_i$ | 0.25 | 0.75 | 1.25 | 1.75 | 2.25 |
|---|---|---|---|---|---|
| $y_i$ | 0.28 | 0.57 | 0.68 | 0.74 | 0.79 |

   dopasować model $y = a_0(1 - e^{-a_1 x})$.

**Odp.:** $\hat{a}_0 = 0.7286$, $\hat{a}_1 = 1.5019$.

2. Skoczek skacze ze spadochronem z balonu napełnionego gorącym powietrzem. Tabela przedstawia jego prędkości [m/s] w kolejnych chwilach czasu [s]:

| $t_i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_i$ | 10 | 16.3 | 23 | 27.5 | 31 | 35.6 | 39 | 41.5 | 42.9 | 45 | 46 | 45.5 | 46 | 49 | 50 |

Dopasować do tych danych modele

$$v(t) = \frac{gm}{c}\left(1 - e^{(-c/m)t}\right)$$

oraz

$$v(t) = \frac{gm}{c}\left(\frac{t}{3.75 + t}\right)$$

gdzie: $g$ – stała grawitacji (9.8 m/s$^2$), $m$ – masa skoczka (68.1 kg), $c$ – współczynnik oporu powietrza (12.5 kg/s). Który z tych modeli jest dokładniejszy?

3. Spektroskop zmierzył intensywność odbieranego sygnału dla różnych długości fali. Wyniki przedstawia tabela:

| $x_i$ | $y_i$ | $x_i$ | $y_i$ |
|---|---|---|---|
| 0 | 25059 | 8 | 295849 |
| 1 | 34459 | 9 | 273272 |
| 2 | 56923 | 10 | 225068 |
| 3 | 109885 | 11 | 171780 |
| 4 | 152544 | 12 | 126180 |
| 5 | 198619 | 13 | 70684 |
| 6 | 256505 | 14 | 43297 |
| 7 | 289850 | 15 | 25515 |

Dopasować do tych danych model

$$y = k_1 e^{k_2(x-k_3)^2}$$

(a) doprowadzając sytuację do przypadku regresji liniowej, (b) stosując bezpośrednio metody regresji nieliniowej. Porównać jakość dopasowania z zastososwaniem wykresu. Jakiej dlugości fali odpowiada maksimum sygnału?

Zadanie powtórzyć dla danych

| $x_i$ | $y_i$ | $x_i$ | $y_i$ |
|---|---|---|---|
| 0 | 569322 | 8 | 150058 |
| 1 | 647595 | 9 | 81671 |
| 2 | 871287 | 10 | 57118 |
| 3 | 904318 | 11 | 32746 |
| 4 | 820139 | 12 | 25717 |
| 5 | 700099 | 13 | 17639 |
| 6 | 434252 | 14 | 13063 |
| 7 | 216687 | 15 | 11527 |

Wytłumaczyć istotną różnice między rezultatami otrzymanymi w punktach (a) i (b).