

Regresja liniowa z zakłóconymi zmiennymi objaśniającymi

Krystyna Maciąg

Uniwersytet Marii Curie Skłodowskiej

Czesław Stępnia

Uniwersytet Rzeszowski

Podstawowym celem regresji liniowej jest prognozowanie wartości zmiennej objaśnianej y na podstawie zaobserwowanych wartości zmiennych objaśniających x_1, \dots, x_k . W klasycznym modelu zakłada się, że zmienne objaśniające są kontrolowane, czyli deterministyczne, a zmienna y jest obserwowana z pewnym błędem, czyli jej obserwacja jest pewną zmienną losową (Y). Powszechnie stosowaną metodą prognozowania wartości y jest *zwykła metoda najmniejszych kwadratów*, polegająca na minimalizacji sumy kwadratów odchyleń wektora obserwacji $Y = (Y_1, \dots, Y_n)$ od proponowanej prognozy liniowej $\hat{Y} = (\sum_{j=1}^k \alpha_j x_{1j}, \dots, \sum_{j=1}^k \alpha_j x_{nj})$.

W ostatniej dekadzie pojawiło się wiele artykułów i kilka książek traktujących zmienne objaśniające jako zmienne losowe powstałe z zakłócenia pewnych nieobserwowalnych wartości deterministycznych. W takim modelu można rozważać prognozowanie nie tylko nieznanymi wartościami zmiennych objaśnianych, ale także (nieobserwowalnych) wartości zmiennych objaśniających. Problem o którym mowa nosi nazwę *totalnie najmniejszych kwadratów* i - jak dotąd - jest głównie domeną numerycznej algebry liniowej. Co ciekawe, jego rozwiązania nie są funkcjami liniowymi obserwacji.

Naszym celem jest elementarne wprowadzenie w tę tematykę.