Legkisebb négyzetek módszere

Tekintsünk egy időben változó q fizikai mennyiséget, melyen m különböző időpillanatban mérést hajtunk végre. Az így kapott adathalmaz:

$$\{(t_1, q(t_1)), (t_2, q(t_2)), \dots, (t_m, q(t_m))\}.$$

Célunk, hogy a mért adataink alapján a lehető legjobban leírjuk a vizsgált fizikai mennyiség időbeli változását. Ezt úgy akarjuk megtenni, hogy a kapott adatpontjainkra egy (legfeljebb) $(n-1)$-ed fokú ($n \leq m$) közelítő polinomot illesztünk:

$$p(t) = x_0 + x_1 t + x_2 t^2 + \cdots + x_{n-1} t^{n-1}.$$

Az ebben szereplő $x_0, x_1, \dots, x_{n-1}$ ismeretlen együtthatókat az ún. legkisebb négyzetek módszerében úgy választjuk meg, hogy a polinom mért értéktől való $\varepsilon_i = |p(t_i) - q(t_i)|$, $(i = 1, 2, \dots, m)$ eltéréseinek, azaz a közelítés hibáinak

$$F(x_0, x_1, \dots, x_{n-1}) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (p(t_i) - q(t_i))^2 = (A\vec{x} - \vec{b})^T (A\vec{x} - \vec{b}).$$

négyzetösszege minimális legyen, ahol

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^{n-1} \end{pmatrix}_{m \times n}, \quad \vec{x} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}, \text{ és } \vec{b} = \begin{pmatrix} q(t_1) \\ q(t_2) \\ \vdots \\ q(t_m) \end{pmatrix}.$$

Tehát a többváltozós $F(x_0, x_1, \dots, x_{n-1})$ függvény minimumát keressük. Egy többváltozós függvény lokális szélsőértéke létezésének szükséges feltételét adja az alábbi tétel.

1. Tétel. *Ha az $f\colon \mathbb{R}^n \to \mathbb{R}$ függvény az $\vec{x}$ pontban minden változója szerint parciálisan differenciálható, és f-nek $\vec{x}$-ben lokális szélsőértéke van, akkor f-nek az $\vec{x}$ pontban minden parciális deriváltja* 0, *azaz* $\frac{\partial f}{\partial x_j}(\vec{x}) = 0$.

Azaz, az $F(x_0, x_1, \dots, x_{n-1})$ függvényünk esetleges szélsőértékének megtalálásához a

$$\frac{\partial F}{\partial x_j}(x_0, x_1, \dots, x_{n-1}) = \sum_{i=1}^{m} 2(p(t_i) - q(t_i))\Big(\frac{\partial p}{\partial x_j}(t_i)\Big) = 0 \quad (j = 0, 1, \dots, n-1)$$

egyenleteket kell megoldanunk. Átrendezve és felhasználva a $p(t)$ polinom alakját, illetve elvégezve a parciális deriválást, $\frac{\partial p}{\partial x_j}(t_i) = t_i^j$:

$$\sum_{i=1}^{m} t_i^j \sum_{k=0}^{n} x_k t_i^k = \sum_{k=0}^{n} x_k \sum_{i=1}^{m} t_i^{j+k} = \sum_{i=1}^{m} q(t_i) t_i^j \quad (j = 0, 1, \dots, n-1).$$

Ezzel az x_k-kra egy lineáris egyenletrendszert kaptunk n egyenlettel és n ismeretlennel, melyet a korábban bevezetett jelöléssel

$$A^T A\vec{x} = A^T\vec{b}$$

alakban írhatunk. Mivel az A mátrix egy olyan Vandermonde-mátrix, melynek oszlopai lineárisan függetlenek (hisz $t_i \neq t_j$, ha $i \neq j$, mert mérési időpillanatok), így rangja n, ezért $A^T A$ invertálható. Innen, az egyértelmű megoldás:

$$\vec{x} = (A^T A)^{-1} A^T \vec{b} := \vec{c}$$

A többváltozós függvények kalkulusa ismeretében (F második deriváltja segítségével) nem nehéz belátni, hogy az így kapott $\vec{c}$-vel adódó $F(\vec{c})$ szélsőérték minimum. Továbbá, mivel $|\vec{x}| \to \infty$ esetén $F(\vec{x}) \to \infty$, így az egyetlen lokális minimum egyben globális minimum is.
Érdemes megemlíteni, hogy $m = n$ esetén $\vec{c} = A^{-1}\vec{b}$, így $F(\vec{c}) = 0$, vagyis a $p(t)$ polinom felveszi a mért $q(t_i)$ értékeket. Ilyenkor az ismertetett eljárást a matematikában Lagrange-interpolációnak nevezik.