5. Funktsioonide lähendamine vähimruutude meetodiga

Olgu sõlmedes

$$a = x_0 < x_1 < \ldots < x_n = b$$

antud funktsiooni väärtused $f(x_0), \ldots f(x_n)$. Otsime mingisse etteantud funktsioonide klassi (nt polünoomid, splainid) kuuluvat funktsiooni $\Phi(x)$, mis lähendaks funktsiooni $f(x)$ nendes sõlmedes. Kõige loomulikum on sellisel juhul lahendada interpolatsiooniülesanne, st määrata $\Phi(x)$ nii, et $\Phi(x_i) = f(x_i)$, $i = 0, \ldots, n$. Paraku ei ole alati võimalik interpolatsiooniülesannet lahendada. Näiteks, kui me otsime $\Phi(x)$ polünoomide hulgast ja sõlmi on rohkem kui $\Phi(x)$ kordajaid, siis on interpolatsiooniülesanne ülemääratud ja seega ei tarvitse tal lahendit olla.

Püstitame järgmise üldisema ülesande: leida $\Phi(x)$ nii, et ta langeks sõlmedes $x_0, \ldots, x_n$ "kõige paremini" kokku funktsiooni $f(x)$ väärtustega. Funktsioonide $\Phi(x)$ ja $f(x)$ erinevuse mõõduks sõlmedes võtame järgmise vahede ruutude summa:

$$J(\Phi) = \sum_{i=0}^{n} \varkappa_i \left(\Phi(x_i) - f(x_i)\right)^2 , \tag{8.1}$$

kus $\varkappa_i > 0$. Otsime niisugust funktsiooni $\Phi(x)$, mille korral on $J(\Phi)$ minimaalne. See on nn *vähimruutude ülesanne*. Seejuures nimetatakse minimeeritavat suurust $J(\Phi)$ *vähimruutude funktsionaaliks*. Suurused $\varkappa_i$ on *kaalud*. Kaalud võimaldavad varieerida funktsioonide $\Phi(x)$ ja $f(x)$ kokkulangemise määra erinevates sõlmedes. Vähimruutude meetodit nimetatakse ka *regressiooniks*.

Kui ei ole põhjust eelistada osasid sõlmi teistele, siis võib kaalud võtta võrdseks ühega ning $J(\Phi)$ on järgmisel kujul:

$$J(\Phi) = \sum_{i=0}^{n} \left(\Phi(x_i) - f(x_i)\right)^2 .$$

Üldjuhul ei läbi vähimruutude meetodiga saadava lähendi Φ graafik etteantud punkte $P(x_i, f(x_i))$, vaid jookseb nende lähedalt läbi.

Kuna $J(\Phi)$ on mittenegatiivsete arvude summa, ei saa ta omada nullist väiksemaid väärtusi, st alati $J(\Phi) \geq 0$. Märgime, et kui funktsionaali $J(\Phi)$ miinimum võrdub nulliga, siis on see miinimum ühtlasi interpolatsiooniülesande

$$\Phi(x_i) = f(x_i) , \quad i = 0, \ldots, n, \tag{8.2}$$

lahendiks. Tõepoolest, kui $J(\Phi) = 0$, siis $\Phi(x_i) - f(x_i) = 0$ iga $i = 0, \ldots, n$ korral, millest tuleneb (8.2). Kui interpolatsiooniülesandel lahendit ei ole, siis on $J(\Phi)$ miinimum nullist suurem.

Juhul kui Φ on ühtlasi ka interpolatsiooniülesande lahend, siis läbib Φ graafik kõiki punkte $P(x_i, f(x_i))$, $i = 0, \ldots, n$.

Mõned praktikas enam esinevad erijuhud.

Lineaarne vähimruutude meetod ehk *lineaarne regressioon.* Lähendit otsitakse järgmisel kujul:

$$\Phi_1(x) = c_1 x + c_2.$$

Mittelineaarne vähimruutude meetod ehk *mittelineaarne regressioon.* Lähendit otsitakse järgmisel kujul:

$$\Phi_m(x) = c_1 x^m + c_2 x^{m-1} + \ldots + c_m x + c_{m+1},$$

kus $m \geq 2$. Selle alamjuhud on omakorda *ruutregressioon* ($m = 2$), *kuupregressioon* ($m = 3$) jne.

Eksponentsiaalne vähimruutude meetod ehk *eksponentsiaalne regressioon.* Lähendit otsitakse järgmisel kujul:

$$\Phi(x) = ae^{bx}.$$

Eksponentsiaalselt regressioonilt saab väga lihtsalt üle minna lineaarsele. Selleks tuleb ülesandes antud funktsiooni väärtusi ja otsitavat lähendit logaritmida. Nimelt $\Phi(x)$ logaritm on lineaarne funktsioon:

$$\Phi_1(x) = \ln \Phi(x) = \ln \left(ae^{bx}\right) = \ln a + bx = c_1 x + c_2, \quad \text{kus } c_1 = b,\ c_2 = \ln a.$$

Kui arvutada tabelis antud funktsiooni väärtuste logaritmid $f_1(x_i) = \ln f(x_i)$ ja lahendada lineaarne vähimruutude ülesanne logaritmitud tabeliga, siis saadava lineaarse lähendi $\Phi_1(x) = c_1 x + c_2$ kordajate kaudu on võimalik arvutada ka esialgsele tabelile vastava eksponentsiaalse lähendi kordajad:

$$a = e^{c_2}\ , \ \ b = c_1.$$