Kovács József mérnök százados
főiskolai adjunktus
Repülő Szakág Tanszék, oktató

ISMERETLEN JELLEMZŐK ÉRTÉKÉNEK BECSLÉSE A LEGKISEBB NÉGYZETEK MÓDSZERÉVEL

A cikkben a szerző a matematikai statisztikából ismert becslés feladatát közvetett mérések esetére alkalmazza. A feladatot a legkisebb négyzetek módszerével oldja meg, determinisztikus és statisztikus megközelítésben. A determinisztikus megközelítés arra az esetre vonatkozik, amikor a mért és a meghatározni kívánt jellemzők közötti kapcsolat nem pontosan ismert, míg a statisztikus megközelítés a véletlen mérési hibákat és az egyes mérések eltérő mérési pontosságát veszi figyelembe.

Sok gyakorlati esetben van szükség ismeretlen jellemzők összességének meghatározására olyan mérések eredményei alapján, amelyek általában hibákat tartalmaznak. Az ismeretlen jellemzők így meghatározott közelítő értékeit becsült értéknek vagy becslésnek, a becslés meghatározásának folyamatát pedig becslési feladatnak nevezzük. A becslési feladat megoldása a mérési eredmények és adatok feldolgozásának egyik formája [1].

Legyen X_1, $X_2,\ldots,X_n$ – ismeretlen jellemzők összessége, amelyek becslésének meghatározására van szükség. Magukat a becsült értékeket jelöljük x_1, $x_2,\ldots,x_n$. Műszaki feladatok megoldása során gyakran fordulnak elő közvetett mérések, amikor nem közvetlenül X_1, $X_2,\ldots,X_n$ értékeket, hanem azokkal valamilyen függvénykapcsolatban álló, y_1, $y_2,\ldots y_k$ jellemzőket mérünk, vagyis:

$$
\begin{aligned}
y_1 &= f_1\,(X_1\,,\,X_2\,,\ldots,\,X_n)\,,\\
y_2 &= f_2\,(X_1\,,\,X_2\,,\ldots,\,X_n)\,,\\
&\;\;\vdots\\
y_k &= f_k\,(X_1\,,\,X_2\,,\ldots,\,X_n)\,.
\end{aligned}
\qquad (1)
$$

A továbbiakban csak azt az esetet vizsgáljuk, amikor a mért jellemzők az ismeretlen jellemzőkkel lineáris függvénykapcsolatban vannak. Ekkor az (1) egyenlet a következő formában írható fel:

$$\begin{array}{l} h_{11}X_1 + h_{12}X_2 + \ldots + h_{1n}X_n = y_1, \\ h_{21}X_1 + h_{22}X_2 + \ldots + h_{2n}X_n = y_2, \\ \quad \vdots \\ h_{k1}X_1 + h_{k2}X_2 + \ldots + h_{kn}X_n = y_k, \end{array} \qquad (2)$$

ahol h_{ij}- ismert együtthatók.

A (2) kifejezéseket mint k lineáris algebrai egyenletből álló, n ismeretlenes egyenletrendszert vizsgálhatjuk, amelyben az ismeretlenek X_j, $j=\overline{1,n}$ értékek lesznek. A továbbiakban azt az esetet vizsgáljuk, amikor $k>n$, vagyis az egyenletek száma meghaladja az ismeretlenek számát, és a rendszer matematikailag nem rendelkezik megoldással (ha létezne pontos, egyértelmű megoldás, akkor a becslésre nem lenne szükség, mert az ismeretlenek a méréssel pontosan meghatározhatók lennének). Az, hogy az egyenletrendszernek matematikailag nincs pontos megoldása, fizikailag a következő okokra vezethető vissza [1]:

- az X_j és y_i közötti fizikai kapcsolatokat csak közelítően írja le az egyenletrendszer (például a h_{ij} együtthatók nem pontosan ismertek),
- y_i, $i=\overline{1,k}$ értékei hibákkal kerülnek meghatározásra.

Akkor az ismeretlen értékek becslésének feladatát vizsgálhatjuk úgy is, mint a (2) egyenletrendszer közelítő megoldásainak meghatározását feltételezve, hogy y_i, $i=\overline{1,k}$ értékeit mérések sorozatával megállapítottuk. A közelítő megoldást tekinthetjük a becslési feladat megoldásának, tehát a becsült értéknek.

Az összes lehetséges közelítő megoldás közül célszerű a valamilyen értelemben legjobb megoldás kiválasztása. Ilyen megoldás megkeresésének egyik útja a legkisebb négyzetek módszerének alkalmazása (a módszer bemutatása a [2], [3] és [4] irodalmakban is megtalálható). A módszeren belül megkülönböztethetünk determinisztikus és statisztikus megközelítést. A determinisztikus megközelítést akkor alkalmazhatjuk, ha véletlen mérési hibák nincsenek, vagy az ilyen hibák statisztikai jellemzőit nem ismerjük. A statisztikus megközelítésnél a véletlen hibák statisztikai jellemzőit ismerteknek tételezzük fel. Az adatfeldolgozáshoz használt matematikai apparátus ebben az esetben lehetővé teszi a mérések eltérő pontosságának figyelembe vételét és a kapott eredmények pontosságának értékelését is [1].

A (2) egyenletrendszernek a legkisebb négyzetek módszerével meghatározott közelítő megoldásait jelöljük a fentieknek megfelelően x_1, x_2, ..., x_n. Mivel ezek a megoldások közelítőek, így behelyettesítésük az eredeti egyenletbe nem eredményezi az egyenlőség teljesülését. X_j értékeinek x_j értékekkel való helyettesítése után ahhoz hogy az egyenlőségek valóban teljesüljenek az egyenletek bal oldalát ki kell egészíteni valamilyen b_i, $i=\overline{1,k}$ értékekkel. Így az egyenletrendszerünk most:

$$\begin{array}{l} h_{11}x_1 + h_{12}x_2 + \dots + h_{1n}x_n + b_1 = y_1, \\ h_{21}x_1 + h_{22}x_2 + \dots + h_{2n}x_n + b_2 = y_2, \\ \quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad \vdots \\ h_{k1}x_1 + h_{k2}x_2 + \dots + h_{kn}x_n + b_k = y_k. \end{array} \tag{3}$$

Vezessük be a megoldás minőségi jellemzőjét a b_i "egyeztetlenségi" érték skaláris függvényeként:

$$J = f(b_1\ , b_2\ , \dots\ , b_k).$$

A legkisebb négyzetek módszerében olyan négyzetes minőségi jellemzőt használunk, amely a b értékek négyzetösszege [1]:

$$J_1 = \sum_{i=1}^{k} b_i^2 = \sum_{i=1}^{k} [y_i - \sum_{j=1}^{n} h_{ij} x_j]^2. \quad (4)$$

Az eredeti egyenletrendszer optimális megoldása a legkisebb négyzetek módszerével olyan x_j, $j=\overline{1,n}$ érték, amely biztosítja a J_1 minőségi jellemző minimális értékét, vagyis a (4) négyzetösszegek minimumát.

$$J_1 = \sum_{i=1}^{k} [y_i - \sum_{j=1}^{n} h_{ij} x_j]^2 = \min. \quad (5)$$

Ebből a kritériumból ered a módszer elnevezése is. Mivel a J_1 függvényt minimalizálni kell, így azt költségfüggvénynek is tekinthetjük [2],[3].

Mivel a minőségi jellemző az x_j, $j=\overline{1,n}$ becsült értékek függvénye, így az (5) kritérium teljesülésének feltétele a következő egyenletrendszerrel írható le:

$$\delta J_1 / \delta x_j = 0 \ , \ j = \overline{1,n}. \quad (6)$$

Ezek az egyenletek szolgálnak alapul a keresett x_j becslések meghatározásához. A parciális deriváltakat kifejtve:

$$\frac{\delta J_1}{\delta x_j} = -2 \sum_{i=1}^{k} h_{ij} y_i + 2 \sum_{i=1}^{k} h_{ij} \sum_{j=1}^{n} h_{ij} x_j \ ,$$

és $j = \overline{1,n}$ esetére sorban egyenlővé téve zérussal a következő egyenletrendszerhez jutunk:

$$\begin{array}{l} \sum_{i=1}^{k} h_{i1} \sum_{j=1}^{n} h_{ij} x_j = \sum_{i=1}^{k} h_{i1} y_i \ , \\ \sum_{i=1}^{k} h_{i2} \sum_{j=1}^{n} h_{ij} x_j = \sum_{i=1}^{k} h_{i2} y_i \ , \\ \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \sum_{i=1}^{k} h_{in} \sum_{j=1}^{n} h_{ij} x_j = \sum_{i=1}^{k} h_{in} y_i \ . \end{array} \quad (7)$$

A (7) egyenletrendszer n egyenletből áll és n ismeretlent tartalmaz. Bizonyítható, hogy a rendszer egyenleteinek együtthatóival felírt determináns nem zérus értékű, tehát a (6) egyenletrendszer $x_1, x_2, \ldots, x_n$ változókra egyetlen megoldást ad. Ez a megoldás egyidejűleg az eredeti (2) egyenletrendszer megoldása a legkisebb négyzetek módszerével, determinisztikus megközelítésben.

Legyen adott az ismeretlen X_j, $j=\overline{1,n}$, és a mért y_i, $i=\overline{1,k}$, értékek közötti összefüggés továbbra is a (2) egyenletrendszerrel, ahol $k \geq n$. Az előző esettől eltérően az egyenletrendszer most rendelkezzen zérustól eltérő értékű, pontos megoldással; de y_i értékeket hibákkal határozzuk meg, tehát a mérések eredményei nem y_i, hanem azoktól eltérő, de értékekben közeli jellemzők:

$$z_i = y_i + \nu_i \ , \quad i=\overline{1,k} \ , \qquad (8)$$

ahol ν - véletlen mérési hibák.

Mivel ν_i véletlen értékek, így z_i szintén véletlen értékű lesz.

Az elmondottak figyelembevételével a (2) egyenletrendszert most így írhatjuk fel:

$$\begin{aligned} h_{11}X_1 + h_{12}X_2 + \ldots + h_{1n}X_n &= z_1, \\ h_{21}X_1 + h_{22}X_2 + \ldots + h_{2n}X_n &= z_2, \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots \\ h_{k1}X_1 + h_{k2}X_2 + \ldots + h_{kn}X_n &= z_k. \end{aligned} \qquad (9)$$

A (9) egyenletrendszerben $k \geq n$ és az egyenletrendszernek nincs pontos megoldása. Ha a z_i véletlen értékek statisztikai jellemzői nem ismertek, akkor a (9) egyenletrendszer megoldható z_i értékeire a legkisebb négyzetek módszerével determinisztikus megközelítésben is, a (7) egyenletek segítségével. Ha a statisztikai jellemzők ismertek, akkor a statisztikus megközelítést kell alkalmaznunk. Ennek segítségével olyan pontosabb megoldást kapunk, amely figyelembe veszi a mérések eltérő pontosságát is [1].

Legyen a ν_i véletlen mérési hibák várható értéke zérus:

$$M[\nu_i] = 0, \quad i=\overline{1,k};$$

szórásnégyzetük pedig általános esetben eltérő:

$$D[\nu_i] = M[\nu_i^2] = \sigma_i^2, \quad i=\overline{1,k};$$

és legyenek a hibák egymástól függetlenek, vagyis:

$$M[\nu_i \nu_j] = 0, \text{ ha } i \neq j \text{ és } i,j=\overline{1,k}.$$

Fogadjuk el továbbá, hogy a véletlen értékek normál eloszlásúak:

$$f_{\nu_i}(\alpha) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{\alpha^2}{2\sigma_i^2}\right), \quad i=\overline{1,k}.$$

Ennek megfelelően a z_i véletlen értékekre is felírhatjuk:

$$M[z_i] = M[y_i + \nu_i] = y_i,$$

$$D[z_i] = D[\nu_i] = \sigma_i^2, \qquad (10)$$

$$f_{z_i}(\beta_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{(\beta_i - y_i)^2}{2\sigma_i^2}\right).$$

Az x_j becslések meghatározásához a (2) egyenleteket felhasználva helyettesítsük az $f_{z_i}(\beta_i)$ kifejezésében y_i értékeket mint $\sum_{j=1}^{n} h_{ij} x_j$:

$$f_{z_i}(\beta_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\frac{1}{2\sigma_i^2}\left(\beta_i - \sum_{j=1}^{n} h_{ij} X_j\right)^2\right].$$

Legyen A_i véletlen esemény z_i véletlen értékeknek a β_i-től $\beta_i + d\beta_i$-ig terjedő intervallumba esése, amelynek valószínűsége:

$$P[A_i] = P[\beta_i \le z_i \le \beta_i + d\beta_i] = f_{z_i}(\beta_i)\, d\beta_i.$$

Legyen továbbá A véletlen esemény valamennyi z_i, $i=\overline{1,k}$ értékeknek a megfelelő $\beta_i \div (\beta_i + d\beta_i)$ intervallumba esése, vagyis $A = A_1 A_2 \ldots A_k$. Az elfogadott kikötések értelmében z_i véletlen értékek egymástól függetlenek, ezért A_i események szintén egymástól függetlenek lesznek. Akkor:

$$P[A] = P[A_1]\, P[A_2] \ldots P[A_k] =$$
$$= f_{z_1}(\beta_1)\, f_{z_2}(\beta_2) \ldots f_{z_k}(\beta_k)\, d\beta_1\, d\beta_2 \ldots d\beta_k =$$
$$= f_{z_1,z_2,\ldots z_k}(\beta_1,\beta_2,\ldots\beta_k)\, d\beta_1\, d\beta_2 \ldots d\beta_k. \qquad (11)$$

ahol

$$f_{z_1,z_2,\ldots z_k}=(2\pi)^{-k/2}\prod_{i=1}^{k}\sigma_i^{-1}\exp\left[-\sum_{i=1}^{k}\frac{1}{2\sigma_i^2}(\beta_i-\sum_{j=1}^{n}h_{ij}X_j)^2\right]$$

Ha β_1, β_2, ..., β_k értékeit rögzítjük, akkor $f_{z_1,z_2,\ldots,z_k}$ függvény $L(X_1,X_2,\ldots,X_n;\beta_1,\beta_2,\ldots,\beta_k)$ likelihood-függvénnyé válik, ami lehetővé teszi $x_1,x_2,\ldots,x_n$ becslések meghatározásában a maximum-likelihood-módszer alkalmazását [2], [3].

Végezzünk véletlen próbát, amely kimeneteleinek valószínűsége nem ismert. A próbák során a nagyobb valószínűségű eredmények gyakrabban fognak előfordulni mint azok, amelyek kisebb valószínűséggel rendelkeznek, ezért az eredmény valószínűségének értékelésekor egyszeri próba esetén a bekövetkező eredmény "a priori" valószínűségét maximálisnak lehet venni.

A maximum-likelihood-módszert a vizsgált feladatra alkalmazva, tegyük fel, hogy P [A]= max, vagyis A esemény valószínűsége maximális. A (11) kifejezésből látszik, hogy ez a követelmény akkor teljesül, ha $X_1=x_1$, $X_2=x_2$, ..., $X_n=x_n$, amely esetben biztosítva van a likelihood-függvény maximuma:

$$L(X_1,X_2,\ldots,X_n;\beta_1,\beta_2,\ldots,\beta_k)=$$
$$=\text{const}\exp\left[-\sum_{i=1}^{k}\frac{1}{\sigma_i^2}(\beta_i-\sum_{j=1}^{n}h_{ij}X_j)^2\right]\quad(12)$$

A felírt kifejezésből következik, hogy a likelihood-függvény maximuma a hatványkitevőben álló kifejezés abszolút értékének minimuma esetén következik be. Ebből írhatjuk fel a (9) egyenletrendszer $x_1,x_2,\ldots x_n$ megoldásainak optimumkritériumát a következő alakban:

$$J_2=\sum_{i=1}^{k}\frac{1}{\sigma_i^2}(z_i-\sum_{j=1}^{n}h_{ij}x_j)^2=\min.\quad(13)$$

A kritériumból következő egyenletrendszerünk ekkor:

$$\frac{\delta J_2(x_j)}{\delta x_j}=0\ ,\ j=\overline{1,n}\qquad(14)$$

A parciális deriváltakat meghatározva:

$$\frac{\delta\, J_2(x_j)}{\delta\,(x_j)} = -2\sum_{i=1}^{k}\frac{1}{\sigma_i^2}h_{ij}z_i + 2\sum_{i=1}^{k}\frac{1}{\sigma_i^2}h_{ij}\sum_{j=1}^{n}h_{ij}x_j.$$

Ezt a kifejezést $j=1, 2, \ldots, n$ értékekre sorban egyenlővé téve zérussal a következő egyenletrendszert írhatjuk fel:

$$\sum_{i=1}^{k}\frac{1}{\sigma_i^2}h_{ij}\sum_{j=1}^{n}h_{ij}x_j = \sum_{i=1}^{k}\frac{1}{\sigma_i^2}h_{ij}z_i;\quad j=\overline{1,n}. \qquad (15)$$

A (15) egyenletrendszer megoldása egyidejűleg a vizsgált (9) rendszernek a legkisebb négyzetek módszerével meghatározott megoldása is.

A (13), (5) optimumkritériumokat és a (15), (7) egyenleteket összehasonlítva látjuk, hogy statisztikus megközelítésnél az y_i értékek mérési pontosságát a mérési hibák σ_i^2 szórásnégyzetével vesszük figyelembe. Azonos pontosságú méréseknél $\sigma_i=\sigma$ és a (13), (15) egyenletek megegyeznek az (5), (7) egyenletekkel. A (15) egyenleteket értelmezhetjük lineáris kapcsolatok rendszereként a $z_1, z_2, \ldots, z_k$ mért értékek és az $X_1, X_2, \ldots, X_n$ jellemzők között.

Mivel z_i, $i=\overline{1,k}$ véletlen értékek, így az x_j, $j=\overline{1,n}$ becslések is véletlen jellemzők lesznek. A becslések statisztikai jellemzői a következők [1]:

- az x_j becslések torzítatlanok az ismeretlen X_j értékekhez képest, vagyis

$$M[x_j] = X_j,\ j=\overline{1,n}\ ;$$

- a becslések hibáinak szórásnégyzete minimális. A becslés hibája alatt itt az $x_j - X_j$, $j=\overline{1,n}$ különbséget értjük;
- ha a mérési hibák normál eloszlásúak, akkor a becslések is normál eloszlásúak lesznek.

Gyakran nem a fő $X_1, X_2, \ldots, X_n$ jellemzőket, hanem azok lineáris függvényét kell becsülnünk:

$$S = a_1X_1 + a_2X_2 + \ldots + a_nX_n. \qquad (16)$$

Ezt a függvényt lineáris alaknak, S értékének becslését pedig a fő paraméter lineáris alakja becslésének hívjuk. Az S jellemző optimális becslése ebben az esetben:

$$s = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

ahol x_1, x_2, ... ,x_n - X_1, X_2, ...,X_n értékeknek a legkisebb négyzetek módszerével meghatározott optimális becslései. Az s becslés optimalitása abban áll, hogy az torzítatlan, hibája pedig minimális szórásnégyzettel rendelkezik, vagyis:

$$M\,[s] = S,\ D\,[s - S] = \min.$$

Felhasznált irodalom

[1] - V. G. Taraszov: Obrabotka informacii v avtomatyizírovannih szisztyemah upravlenyija (VVIA im. prof. N.E. Zsukovszkogo, Moszkva, 1974).

[2] - Granino A.Korn-Theresa M.Korn: Matematikai kézikönyv műszakiaknak (Műszaki Könyvkiadó, Bp., 1975).

[3] - I.N. Bronstejn-K.A. Szemengyajev: Matematikai zsebkönyv (Műszaki Könyvkiadó, Bp., 1987).

[4] - Edwin F. Beckenbach: Modern matematika mérnököknek II. kötet (Műszaki Könyvkiadó, Bp., 1965).

[illegible]

[illegible]

[illegible]