

Szegedi Orvostudományi Egyetem Számítástechnikai
Központ és Semmelweis Orvostudományi Egyetem Szá-
mitástechnikai Csoport

Nemlineáris paraméterbecslés dinamikus biológiai rendszerek
modellezése során - I: paraméterbecslő eljárások és a modell
-illeszkedés statisztikai elemzése

Eller József, Győri István és Magyar Gábor

Bevezetés

Matematikai modellek alkalmazása az orvosi biológiában két fontos célt szolgálhat. Az egyik hipotetikus modellek konstruálása és tesztelése megfigyelt adatokhoz való illesztés útján. Amennyiben egy modell jól magyarázza a mérési adatokat, jól "illeszkedik" hozzájuk, úgy elfogadható az illető biológiai jelenség leírására. Másik fontos cél lehet már elfogadott, jól bevált modellek biológiailag fontos paramétereinek becslése a megfigyelési adatok alapján. Ez utóbbira számos példát szolgáltat pl. a gyógyszerkinetika /lásd (14)/.

Célunk az volt, hogy a biológiában gyakran vizsgált, időben lezajló dinamikus folyamatok matematikai modelljeinek illesztésére, paraméterbecslésére és az illeszkedés és a becsült paraméterek jóságának eldöntésére jól használható számítástechnikai eljárást dolgozzunk ki. E cél érdekében a programot a BMDP-77 statisztikai programrendszer (3) keretein belül implementáltuk, a BMDP3R nemlineáris regressziós program továbbfejlesztéseként. E rendszer vezérlőnyelve lehetővé teszi változatos méretű és paraméterezésű feladatok

egyszerű leírását. A BMDP3R program paraméterezését olyan paramétermegadási lehetőségekkel bővítettük ki, amely lehetővé teszi a különböző dinamikus modellek egyszerű megadását a program számára.

Az időben lejátszódó, dinamikus jelenségek kvantitatív leírása során rendszerint differenciálegyenletekből indulunk ki. Egyszerűbb esetekben a differenciálegyenletek analitikus megoldása is rendelkezésünkre áll. Például ismert (5), hogy a biológiai kompartment-rendszerek leírására leggyakrabban használt lineáris állandóegyütthetős differenciálegyenletek megoldása általában előáll exponenciális tagok összegeként. Azonban az exponenciálisok illesztésével szemben a modell-differenciálegyenletek megoldásainak - a differenciálegyenlet paraméterei szerint való - illesztése a következő előnyökkel jár:

a/ a modell-differenciálegyenletben szereplő, biológiailag jól értelmezhető paraméterek értékére és szórására kapunk becslést, nem pedig ezek nemlineáris függvényeire /pl. a sajátértékekre, melyek több transzportlépés eredőjét képezik/;

b/ figyelembe vehetők a biológiailag értelmezett paraméterekre /pl. transzportegyütthetőkra/ a biológus által kirótt, vagy természetesen fennálló korlátok /pl. nemnegativitás/, ill. feltételek.

Több esetben azonban nem állandó, hanem időben változó együtthetőju lineáris differenciálegyenletek alkalmazására van szükség. Például az *in vitro* kinetikai mérések során a vizsgálati idő alatt a transzportegyütthetők lassan, de észrevehető módon megváltoznak. Vagy pl. hosszabb idejű /egy-két hetes/ vizsgálatok során a beteg anyagcsere paraméterei /pl. reakciósebességek/ - különösen terápia alatt - módosulnak, amit mindenképpen figyelembe kell venni.

Az említett esetekben, és általában nemlineáris modell-differenciálegyenletek esetén a megoldás, és így a modell-függvény, nem fejezhető ki véges számú elemi függvény segítségével. Ugyanez igaz a modell-függvény parciális deriváltjaira, melyek számítására a paraméterbecslési eljárás során van szükség. Utóbbiak újabb differenciálegyenletek - az un. *érzékenységi egyenletek* megoldásai segítségével fejezhetőek ki. A modell-differenciálegyenletek és a hozzájuk tartozó érzékenységi egyenletek megoldására hatékony számítási eljárást ismerttet következő előadásunk (11). Ezen előadás további részében a paraméterbecslés és modell-illeszkedés kérdésével foglalkozunk.

A probléma matematikai megfogalmazása

Egy dinamikus rendszer egy $\underline{x}(t)$ állapotváltozó vektorral írható le. Ha feltesszük, hogy a rendszer állapota időbeli változásának $\dot{\underline{x}}(t)$ sebessége egy meghatározott módon /pl. lineárisan/ függ magától az $\underline{x}(t)$ állapottól, és egy $\underline{\theta}$ paramétervektortól és esetleg a t időtől, akkor a rendszert egy

$$\dot{\underline{x}} = \underline{f}(t, \underline{x}, \underline{\theta}) \quad /1/$$

alakú modell-differenciálegyenlet írja le. Az /1/ egyenlet megoldása az időn kívül a $\underline{\theta}$ paramétervektortól is függ, azaz $\underline{x} = \underline{x}(t, \underline{\theta})$, mivel különböző paraméter-értékekhez különböző megoldások tartozhatnak.

Általában magát az $\underline{x}(t, \underline{\theta})$ állapotvektort nem tudjuk közvetlenül megfigyelni, hanem csak ennek valamilyen

$$y(t, \underline{\theta}) = g(\underline{\theta}, \underline{x}(t, \underline{\theta})) \quad /2/$$

függvényét. Azonban a mérésre rakódó zaj miatt valójában egy

$$\tilde{y}(t) = y(t, \underline{\theta}^0) + \varepsilon(t) \quad /3/$$

függvényt mérünk, ill. figyelünk meg t diszkrét értékeire, ahol $\varepsilon(t)$ jelöli a zajt és $\underline{\theta}^0$ a paramétervektornak az adott konkrét rendszerre jellemző "pontos" értéke a modell szerint. A feladat ezen elméleti paramétervektor-érték becslése az $\tilde{y}(t)$ mérési adatokból.

Matematikai-statisztikai megfontolások alapján a paraméterek ismeretlen értékének becslését a *legkisebb négyzetek elve* alapján szokták megadni, amely a $\hat{\underline{\theta}}$ becslő értékre az

$$S(\hat{\underline{\theta}}) := \sum_t [\tilde{y}(t) - y(t, \hat{\underline{\theta}})]^2 = \min! \quad /4/$$

minimumfeltételt írja elő. Ilyen statisztikai szempont pl. az, hogy az $\varepsilon(t)$ hibák korrelálatlansága, zérus várható értéke, azonos varianciája és normális eloszlása esetén ez a módszer egybeesik a *maximum-likelihood* módszerrel (2), amely több statisztikailag előnyös tulajdonsággal bír.

Paraméterbecslő iterációs módszerek

Bár többváltozós függvények minimalizálására többféle eljárás ismeretes, de a /4/ minimumfeladat megoldására azok az eljárások a leghatékonyabbak, melyek figyelembe veszik a minimalizálandó $S(\underline{\theta})$ függvény négyzetösszeg-strukturáját (2). E módszerek alapja a négyzetreemelendő $\tilde{y}(t) - y(t, \underline{\theta})$ függvény linearizálása az aktuális, mondjuk k -adik közeli-tés körül. Ezáltal a minimalizálandó $S(\underline{\theta})$ függvényt helyettesítjük az

$$S_k(\underline{\theta}) = S_k(\underline{\theta}^k + \underline{\delta}) = \sum_t [\tilde{y}(t) - y(t, \underline{\theta}^k) - \sum_j \frac{\partial y}{\partial \theta_j}(t, \underline{\theta}^k) \cdot \delta_j]^2 \quad /5/$$

kvadratikus függvénnel, melynek minimumhelyét már viszonylag egyszerű meghatározni, pl. a $\nabla S_k = 0$ feltételből adódó ún. Gauss-féle normálegyenletrendszerből:

$$A^T A \underline{\delta} = A^T \underline{r} \quad /6/$$

ahol

$$\left. \begin{aligned} A &= (a_{ij}), \quad a_{ij} = \frac{\partial y}{\partial \theta_j}(t_i, \underline{\theta}^k), \quad j=1, \dots, p \\ \underline{r} &= (r_i), \quad r_i = \tilde{y}(t_i) - y(t_i, \underline{\theta}^k) \end{aligned} \right\}, \quad i=1, \dots, m;$$

továbbá p a $\underline{\theta}$ paramétervektor komponenseinek száma és m a mérési pontok száma.

A klasszikus Gauss-Newton-módszer a /6/ normálegyenletrendszer megoldásaként adódó $\underline{\delta}^\circ$ vektort használja az aktuális közelítés korrekciójaként és az újabb közelítést $\underline{\theta}^{k+1} = \underline{\theta}^k + \underline{\delta}^\circ$ alakban veszi fel. Azonban az eljárás ilyen formájában nem feltétlenül konvergens, még jó kezdőérték esetén sem (9). A divergencia oka lehet a modell-függvény nagymértékű nemlinearitása vagy rossz illeszkedése a mérési adatokhoz. Ilyen esetekben szokásos a $\underline{\delta}^\circ$ korrekciós vektornak egy olyan $\omega < 1$ relaxációs tényezővel való figyelembe vétele, hogy a

$$\underline{\theta}^{k+1} = \underline{\theta}^k + \omega \cdot \underline{\delta}^\circ \quad /7/$$

helyen már kisebb legyen a minimalizálandó $S(\underline{\theta})$ függvény értéke, mint az előző, $\underline{\theta}^k$ közelítésben. A BMDE3R program pl. $\omega = 1$ -ből kiindulva addig felezi ω -t, amíg e feltétel nem teljesül.

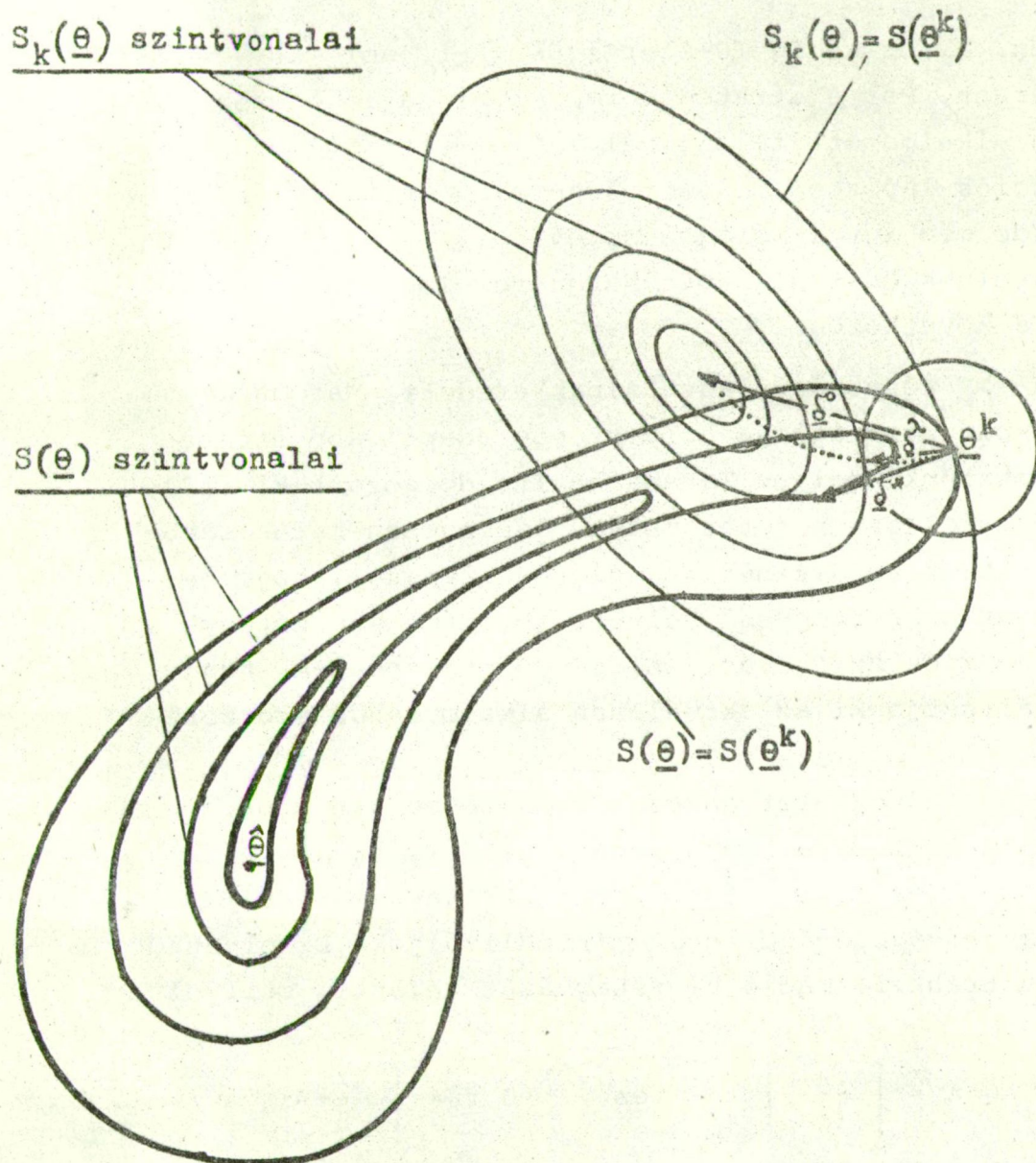
Ha a $\underline{\delta}^\circ$ korrekció alkalmazása nem vezet az $S(\underline{\theta})$ négyzetösszeg csökkenéséhez, annak oka a modell-függvény lineáris közelítésének elégtelensége. Marquardt módszere (2), (9), (12) ilyenkor a korrekciós vektor nagyságának csökkentésével egyidőben annak irányát fokozatosan a minimalizálandó függvény negatív gradiense felé fordítja, melynek irányában lokálisan legmeredekebb a csökkenés. A módszer által definiált $\underline{\delta}^\lambda$ korrekciós vektor az

$$(A^T A + \lambda I) \cdot \underline{\delta}^\lambda = A^T \underline{r} \quad /8/$$

lineáris egyenletrendszer megoldása, ahol $\lambda \geq 0$ valós szám. Ha $\lambda = 0$, akkor a Gauss-Newton-féle $\underline{\delta}^\circ$ korrekciós vektort kapjuk; a λ -t csak akkor növeljük, ha ez szükséges ahhoz, hogy a $\underline{\delta}^\lambda$ korrekció alkalmazásával csökkenjen $S(\underline{\theta})$ értéke. Marquardt bebizonyította, hogy ez elég nagy λ -ra biztosan teljesül, hacsak nem egy stacionárius pontban vagyunk, mert $\|\underline{\delta}^\lambda\|$ λ növekedésével monoton csökken, $\underline{\delta}^\lambda$ iránya pedig egyre jobban egybeesik $S(\underline{\theta})$ negatív gradiensének irányával.

A Marquardt-módszer lokális optimalitási tulajdonsága, hogy $\underline{\delta}^\lambda$ minimalizálja az $S_k(\underline{\theta}^k + \underline{\delta})$ közelítő négyzetösszeget a $\|\underline{\delta}^\lambda\|$ sugaru gömbben /lásd az 1. ábrát/.

Az ábrán kipontozott iv jelzi a lehetséges Marquardt korrekciós vektorok mértani helyét, miközben λ befutja a $[0, \infty)$ intervallumot. A λ paraméter iterációs lépésenkénti megválasztására Marquardt olyan stratégiát javasolt (12), mely szerint az $S(\underline{\theta})$ négyzetösszeg csökkenésének feltétele mellett lehetőleg minél közelebb kell maradni a Gauss-Newton-módszerhez. Ennek megfelelően, ha az előző iterációs lépésben λ egy kis pozitív küszöbszámnál /pl. 10^{-6} / nagyobb volt, akkor próbáljuk meg csökkenteni, pl. a 10-edrésére, amennyiben így is csökken $S(\underline{\theta})$ értéke. Ha



1. ábra

A Marquardt-módszér geometriai szemléltetése a $p = 2$ két-dimenziós esetben. $\underline{\delta}^0$, $\underline{\delta}^\lambda$ és $\underline{\delta}^*$ rendre Gauss-Newton, Marquardt és a legmeredekebb lejtő módszere szerinti korrekciós vektorok, $\hat{\underline{\theta}}$ pedig a keresett minimumhely

nem csökken, akkor addig 10-szerezzük λ -t, amíg egy csökkenést elérünk. Bár e stratégia valamivel hajlékonyabb a BMDP3R-ben alkalmazott felezésnél /mivel itt mindig λ előző iterációs lépésben felvett legutolsó értékéből indulunk ki/, de még így is eléggé merev.

Az általunk összeállított Marquardt-implementáció fő jellemzői a következők.

/1/ A λ tényező választásának eredeti, Marquardt által javasolt stratégiája helyett egy adaptívabb stratégiát alkalmaztunk, melyet Tataba és Ite dolgozott ki (13), és több próbafeladaton letesztelve, lényegesen hatásosabbnak bizonyult az eredetinel. Az eljárás lényege, hogy egyetlen 10-es szorzótényező helyett többféle szorzótényezőt is felvesz és az iteráció előző három lépésének függvényében választja ki az aktuálisan alkalmazandó szorzótényezőt.

/2/ A normálegyenletrendszer képzése és annak pl. Cholesky-féle módszerrel való megoldása helyett az $S_k(\theta)$ linearizálás útján nyert négyzetösszeget *Householder-féle ortogonális triangularizációval* minimalizáljuk. Ez olyan Q ortogonális transzformáció végrehajtását jelenti, mellyel

$$Q A = \begin{bmatrix} -R \\ 0 \end{bmatrix} \quad \text{és} \quad Q \underline{r} = \begin{bmatrix} \underline{c} \\ -\underline{d} \end{bmatrix} \quad /9/$$

teljesül, ahol R p -edrendű négyzetes felső trianguláris mátrix és a Q m -edrendű ortogonális mátrix p számú elemi Householder transzformációs mátrix szorzata. Ekkor a /6/ lineáris egyenletrendszer ekvivalens az

$$R \underline{\delta} = \underline{c} \quad /10/$$

háromszögmátrixu egyenletrendszerrel, amely már egyszerű-

ben oldható meg. A /6/, ill. /10/ egyenletek kondíciós számai közt a

$$\text{cond}(A^T A) = [\text{cond}(R)]^2 \quad /11/$$

összefüggés áll fenn, így az utóbbi módszerrel számolva, a kerekítési hibák felhalmozódása jóval kisebb mértékű, durván szólva annak felel meg, mintha a normálegyenlet-módszer-nél dupla pontosságú aritmetikát használnánk /lásd Lawson-Hanson (10)/. További előnye e megközelítésnek, hogy lehetőség van $\text{cond}(R)$ egyszerű becslésére (10); hátránya viszont, hogy műveletigénye $m(n^2+2n)-n^3/3+O(n^2)$, szemben a normálegyenlet Cholesky módszer $(m/2) \cdot (n^2+3n)+n^3/6+O(n^2)$ műveletigényével. Azonban a mi esetünkben, a differenciálegyenletek numerikus megoldásának lényegesen nagyobb műveletigénye miatt, az említett többlet relative nem sokat tesz ki.

Amennyiben valamely $\lambda > 0$ -hoz tartozó $\underline{\delta}^\lambda$ Marquardt-féle korrekciós vektor meghatározására van szükség, úgy olyan, alkalmasan megválasztott elemi Householder-transzformációk szorzataként előálló Q_λ ortogonális transzformációt határozunk meg, mellyel teljesülnek a

$$Q_\lambda \begin{bmatrix} -R \\ \sqrt{\lambda} I \end{bmatrix} = \begin{bmatrix} R_\lambda \\ 0 \end{bmatrix} \quad \text{és} \quad Q_\lambda \begin{bmatrix} \underline{c} \\ \underline{0} \end{bmatrix} = \begin{bmatrix} \underline{c}^\lambda \\ \underline{d}^\lambda \end{bmatrix} \quad /12/$$

relációk; ekkor a $\underline{\delta}^\lambda$ korrekciós vektor az

$$R_\lambda \underline{\delta}^\lambda = \underline{c}^\lambda \quad /13/$$

háromszögmatrixu lineáris egyenletrendszerből fejezhető ki.

Eddigi tapasztalataink szerinti az így összeállított Marquardt eljárás átlagosan kb. 10-15 százalékkal gyorsabb

a BMDP3R-ben alkalmazott Jannrich-Sampson-féle (8) módosított Gauss-Newton-módszernél. Továbbfejlesztést jelent az is, hogy az általunk implementált programban több mérési görbéhez való együttes illesztés is lehetséges.

Megtartottuk a BMDP3R program által biztosított súlyozási opciót: lehetőség van arra, hogy a program a felhasználó által külsőleg megadott súlyokkal képezett súlyozott hibanégyzetösszeget minimalizáljon. E lehetőség a mérések bizonyos eloszlásai mellett /pl. exponenciális vagy Poisson-eloszlás/ a Jennrich és Moore által leirt módon /lásd (7)/, az iteráció folyamán alkalmasan változtatott súlyozás útján felhasználható az illető eloszlásoknak megfelelő maximum-likelihood becslések meghatározására is.

A modell-illeszkedés statisztikai elemzése

A gyakorlatban az iterációs módszerrel nyert parameterbecslés önmagában nem elégíti ki a modellezőt; szeretné tudni, hogy mennyire megbízható, pontos a becslés, ill., hogy jól illeszkedik-e a modell a mérési adatokhoz. A továbbiakban e két kérdés megválaszolását elősegítő különféle mérőszámokat, ill. mutatókat ismertetünk, melyek egy része használatos a szakirodalomban, egy másik részük használatát pedig mi javasoljuk.

A modell-illeszkedés jóságát nemlineáris regressziós modelleknél általában az átlagos hibanégyzettel és néha az un. nemlineáris korrelációs együtthatóval szokták jellemezni. Az $\tilde{y}_i = \tilde{y}(t_i)$ és $\hat{y}_i = y(t_i, \hat{\theta})$ jelölésekkel, ahol $\hat{\theta}$ a legkisebb négyzetek módszerével nyert becslt érték, az átlagos hibanégyzet /Error Mean Square = EMS/ képlete:

$$EMS = \frac{1}{m-p} \sum_{i=1}^m [\tilde{y}_i - \hat{y}_i]^2 \quad /14/$$

Az átlagos hibanégyzet a mérési hiba varianciájának becslése, négyzetgyöke pedig a hibaszórás becslése.

Az átlagos hibanégyzet hiányossága, hogy önmagában nem jellemzi az illeszkedés megfelelő voltát, csak ha a függvényértékek nagyságrendjéhez viszonyítjuk. Ezért helyette, vagy mellette, javasoljuk a következő relatív mutatószámot, melyet *általánosított variációs együttthatónak* /Generalized Coefficient of Variation = GCV/ nevezhetünk:

$$GCV = \sqrt{\frac{1}{m-p} \sum_{i=1}^m \frac{(\tilde{y}_i - \hat{y}_i)^2}{(\hat{y}_i)^2}} \quad /15/$$

E mutatószám a pontonkénti relatív hibák bizonyos értelemben vett közepe. Az elnevezés oka, hogy a legegyszerűbb, $\tilde{y}_i \sim \theta_1$ modell esetén $\hat{\theta}_1 = \bar{y} := (1/m) \sum_{i=1}^m \tilde{y}_i / GCV$ a jólismert /min-ta-/ variációs együttthatóra redukálódik. E mutató használatának nyilván csak akkor van értelme, ha az illesztett görbe a szóban forgó intervallumon nem válik nullává, ami a gyakorlati esetek nagyobbik részében teljesül.

További dimenziómentes mutatószámok az illeszkedés jóségára a korrelációs típusú mutatók: ezek a lineáris regresszióanalízisből ismert többszörös korrelációs együtttható általánosításával származtathatók. Az interpretációnál azonban vigyázni kell arra, hogy egyetlen így nyert mutató sem valódi korrelációs mérőszám: nincs olyan két valószínűségi változó, melyek korrelációjának becslése lehetne. Az ilyen típusú mutatók közül az un. *nemlineáris korrelációs együtttható* vagy *korrelációs index* (1), (4) használatos a gyakorlatban, melynek képlete:

$$RI = \sqrt{1 - \frac{\sum_{i=1}^m [\tilde{y}_i - \hat{y}_i]^2}{\sum_{i=1}^m [\tilde{y}_i - \bar{y}]^2}} \quad /16/$$

RI értéke annál közelebb van egyhez, minél inkább jobban illeszkednek a mérési adatok a modellfüggvényhez, mint a közönséges átlaghoz.

Egy másik lehetséges általánosítása a lineáris regresszió többszörös korrelációs együtthatójának a mért \tilde{y}_i és a nemlineáris modellből becsült \hat{y}_i értékek közti szokásos /lineáris/ korrelációs együttható, melyet *görbevonalu korrelációs együtthatónak* /CurviLinear correlation coefficient = RCL/ célszerű nevezni:

$$RCL = \frac{\sum_{i=1}^m (\tilde{y}_i - \bar{y}) \cdot (\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^m [\tilde{y}_i - \bar{y}]^2 \sum_{i=1}^m [\hat{y}_i - \bar{\hat{y}}]^2}} \quad /17/$$

Lineáris regressziós modell esetén ismert, hogy $RCL=RI$, nemlineáris modellekben azonban általában $RCL \neq RI$.

Az eddig bemutatott mérőszámok a modell-illeszkedés jóságát olyan értelemben jellemezték, hogy mennyire kicsik az illesztett modellhez tartozó $r_i = \tilde{y}_i - \hat{y}_i$ reziduálok. Ha a reziduálok viszonylag nagyok, annak két fő oka lehet: /a/ a mérési hibák viszonylag nagyok; /b/ a feltételezett regressziós modell nem elég pontosan írja le a vizsgált jelenséget. Hogy az /a/ és /b/ tényezők közül melyik van jelen, azt a *reziduálok analizise* útján döntjük el. Ha a reziduálok nem véletlenszerűen ingadoznak a nulla körül, hanem szisztematikus tényezőt, un. *trendet* tartalmaznak, akkor ez a modell-függvény kisebb vagy nagyobb hiányosságát

tükrözi. Ilyen trendek jelenlétének megállapítása történhet szemrevételezéssel, melyhez jó segítséget jelent a BMDP3R programból átvett rajzoló szubrutin, amely kirajzolja a reziduálokat a becsült \hat{y}_i értékekhez, valamint tetszőleges független változóhoz /tehát pl. az időhöz/ viszonyítva.

A reziduálok véletlenszerűségének kvantitatív jellemzése történhet az

$$RS = \frac{\sum_{i=2}^m (\tilde{y}_i - \hat{y}_i) \cdot (\tilde{y}_{i-1} - \hat{y}_{i-1})}{\sqrt{\sum_{i=2}^m [\tilde{y}_i - \hat{y}_i]^2 \sum_{i=1}^{m-1} [\tilde{y}_i - \hat{y}_i]^2}} \quad /18/$$

sorozatkorrelációs együttható /Serial correlation coefficient/ segítségével. RS nagy /egyhez közeli/ értéke trend jelenlétére utal.

Ezenkívül a reziduálok véletlenszerűségének nullhipotézise tesztelhető az un. *run-szám-próba* (2) segítségével, amely a pozitív és negatív reziduáloknak, ill. a reziduálok jelváltásainak a számán alapul. Ha mind a pozitív, mind a negatív reziduálok száma eléri a 10-et, akkor a próbastatisztika eloszlása már megfelelően közel van a normális eloszláshoz; ekkor a jólismert integráltranszformációval p-statisztikává alakítható; egyébként pedig a megfelelő táblázat használata szükséges a próba elvégzéséhez.

Ami az iterációs módszerrel nyert $\hat{\theta}_i$ paraméterbecslések pontosságát illeti, az jellemezhető például az $M([\hat{\theta}_i - \theta_i^0]^2)$ várható négyzetes hibával; ennek, és egyúttal a $\hat{\theta}_i$ szórásnégyzetének mintából való becslése a linearizált modell és a $\hat{\theta} \sim M(\hat{\theta}) \sim \hat{\theta}$ közelítések alapján:

$$M([\hat{\theta}_i - \theta_i^0]^2) \approx D^2(\hat{\theta}_i) \approx EMS \cdot \hat{c}_{ii}, \quad /19/$$

ahol $\hat{C} = (\hat{c}_{ij}) = [\hat{A}^T \hat{A}]^{-1}$

és \hat{A} a /6/ szerinti parciális deriváltak mátrixa a $\hat{\theta}$ helyen.

A /19/ jobboldalán szereplő mennyiség kis EMS átlagos hibanégyzet mellett is nagy lehet, ha az $\hat{A}^T \hat{A}$ /tapasztalati/ Fisher-féle információs mátrix közel szinguláris; ha szinguláris, akkor végtelen nagy: ez a paraméterek között fennálló függvénykapcsolatot, azaz paraméter-redundanciát jelez. A redundáns paraméterek száma az információs mátrix rangjától függ; ezt legmegbízhatóbban főkomponens-analízis segítségével, a mátrix sajátértékeinek meghatározásán keresztül becsülhetjük.

Programunk meghatározza a Fisher-féle információs mátrix sajátértékeit és sajátvektorait. E sajátértékek alapján a paraméterbecslés egyértelműségének mértékeként a legkisebb és a legnagyobb sajátérték

$$\mu = \lambda_{\min}(\hat{A}^T \hat{A}) / \lambda_{\max}(\hat{A}^T \hat{A}) \quad /20/$$

hányadosát javasoljuk, amely nem más, mint a numerikus analízisből jólismert /cf. (10)/ kondíciós szám reciproka; mindig 0 és 1 közé esik; ha 0, akkor általában nem egyértelműek a paraméterek; ha 1, akkor nagyon jól meghatározottak. Ha μ kicsi, az azt jelenti, hogy a paraméterter nagy kiterjedésű részhalmazán az $S(\hat{\theta})$ négyzetösszeg a minimálishoz igen közeli értéket vesz fel. Ilyen gyengén kondicionált problémáknál gyakran tapasztalható, hogy az iterációs módszer kezdeti "normális" szakasza utáni stádiumban az iterált paraméterközelítések elég sokat "vándorolnak", miközben a minimalizálandó függvény értéke alig csökken.

Végül megjegyezzük még, hogy a $\hat{\theta}$ becsült érték kö-

rül linearizált modellen gyakorlatilag az összes, a többváltozós lineáris regresszióban szokásos próba elvégezhető, ill. konfidencia-intervallumok adhatók; elméleti szempontból azonban ezek eredményei csak megfelelő adag gyanakvással fogadandók el, mivel az említett linearizálás gyakran nem biztosítja a kellő pontosságot.

Programunk a felhasználó informálása céljából valamennyi fentebb ismertetett mutatószámot kinyomtatja. Hogy az alternatív mérőszámok közül a felhasználó melyikre és milyen súllyal támaszkodjék, arra nézve - mivel a nemlineáris regressziós modellek statisztikájának elmélete e téren nem eléggé kidolgozott - nem adható egyértelmű utmutatás; ez a modellezés mindennapi gyakorlatában alakul ki, és függ a vizsgált modell-típusoktól, a mérések minőségétől, valamint attól, hogy melyik mutatószám hogyan válik be a gyakorlatban.

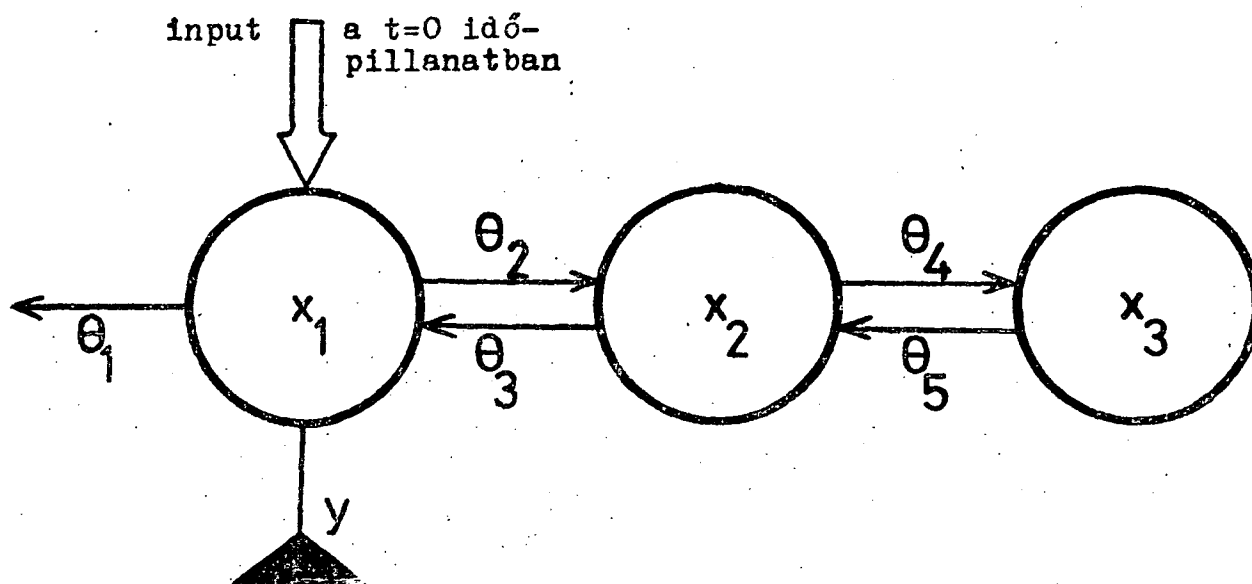
Példa

Szemléltető példa gyanánt a 2. ábrán látható rekeszmodellt (6) választottuk, amely a vérbe /1. rekesz/ beadott radioaktív szulfát felszivódását modellezi. A modellt leíró differenciálegyenletek, ill. a megfelelő kezdeti feltételek a következők:

$$\begin{aligned} \dot{x}_1 &= -(\theta_1 + \theta_2) \cdot x_1 + \theta_3 x_2 & x_1(0) &= 2 \cdot 10^5 \\ \dot{x}_2 &= \theta_1 x_1 - (\theta_3 + \theta_4) \cdot x_2 + \theta_5 x_3 & x_2(0) &= 0 \\ \dot{x}_3 &= \theta_4 x_2 - \theta_5 x_3 & x_3(0) &= 0 \end{aligned}$$

A mérési adatok az ábrának megfelelően az első rekesz megfigyeléséből származnak, tehát

$$y(t, \underline{\theta}) = x_1(t, \underline{\theta})$$

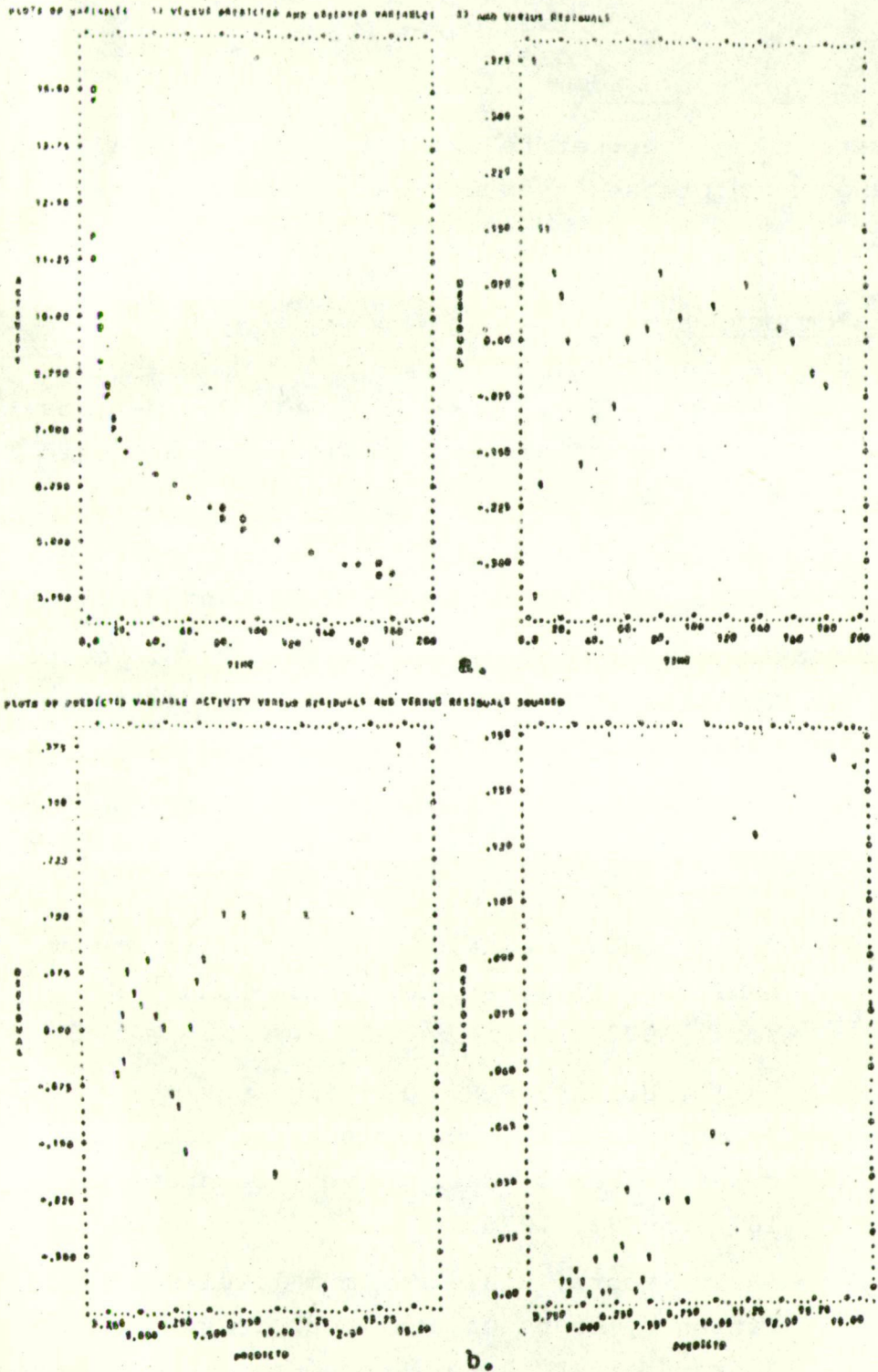


2. ábra

Radioaktív szulfát eloszlásának egy rekeszmodelje

az illesztendő modell-függvény. Az illesztés eredményét a 3. ábra mutatja be.

A 3a. ábrán látható a program által kirajzolt megfigyelt /O=observed/ és becsült /P=predicted/ görbe, valamint a reziduálok görbéje az idő függvényében. Látható, hogy az illeszkedés igen jó, amit a megfelelő mutatók is követnek /RI = 0.9999, GCV = 0.0066/. A reziduálok run-szám-próbája /p = 0.36/, valamint a sorozatkorrelációs együttható /RS = 0.511/ nem mond ellent a reziduálok véletlenszerű voltának, az idő függvényében. Azonban, a 3b. ábrán látható a reziduáloknak a becsült értékekhez viszonyított képén, hogy ezek szórása lineárisan növekvő trendet mutat, ami alapján esetleg érdemes egy, a mérési értékek négyzetének reciprokával súlyozott négyzetösszeg minimalizálását is elvégezni.



3. ábra

A 2. ábrán látható modell illesztésének eredménye

Köszönetnyilvánítás

A szerzők köszönetüket fejezik ki Kányár Bélának jelen munkához nyújtott segítségéért és hasznos megjegyzéséért.

Irodalomjegyzék

- (1) Adam, J., Scharf, J.-H., Enke, H.: Methoden der Statistischen Analyse in Medizin und Biologie. Vlg. Volk und Gesundheit, Berlin, 1977.
- (2) Bard, Y.: Nonlinear Parameter Estimation. Academic Press, New York-San Francisco-London, 1974.
- (3) Dixon, W.J., Brawn, M.B. /szerk./: BMDP-77 Biomedical Computer Programs, P-Series. University of California Press, Los Angeles, 1977.
- (4) Ezekiel, M., Fox, K.A.: Korreláció- és regresszióanalízis. Közgazdasági és Jogi Könyvkiadó, Budapest, 1970.
- (5) Jacquez, J.A.: Compartmental Analysis in Biology and Medicine. Elsevier, Amsterdam-London-New York, 1972.
- (6) Jennrich, R.I., Bright, P.B.: Fitting systems of linear differential equations using computer generated exact derivatives. Technometrics 18: 385-392, 1976.
- (7) Jennrich, R.I., Moore, R.H.: Maximum likelihood estimation by means of nonlinear least squares. Amer. Statist. Assn. Proc. Stat. Comp. Sect., 57-65, 1975.

- (8) Jennrich, R.I., Sampson, P.F.: Application of stepwise regression to non-linear estimation. *Technometrics* 10: 63-72, 1968.
- (9) Kowalik, J., Osborne, M.R.: *Methods for Unconstrained Optimization*. American Elsevier, New York, 1968.
- (10) Lawsen, C.L., Hanson, R.J.: *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, N.J., 1974.
- (11) Magyar G., Galántai A., Kanyár B., Eller J.: Nemlineáris paraméterbecslés dinamikus biológiai rendszerek modellezése során - II. /E kötet köv. közleménye/.
- (12) Marquardt, D.W. I An algorithm for the estimation of non-linear parameters. *SIAM Journal* 8: 181-217, 1963.
- (13) Tataba, T., Ite, R.: Effective treatment of the interpolation factor in Marquardt's nonlinear least-squares fit algorithm. *Computer Journal* 18: 250-251, 1975.
- (14) Wagner, J.G.: *Fundamentals of Clinical Pharmacokinetics*, Drug Intelligence Publ., Hamilton, Illinois, 1975.