

MATEMATIKAI ALGORITMUSOK BIOLÓGIAI ALKALMAZÁSÁBAN ELKÖVETETT HIBÁINK

DR. BÁN ISTVÁN

Örvendetes hétköznapi jelenség ma már, hogy számítógépes munkáink tekintélyes részét teszi ki a különböző matematikai algoritmusok biológiai alkalmazása. Ez a munka sajátos ötvözetét jelenti a számítástechnika, matematika, biológia együttes alkalmazásának. A biológiai feladatok pedig általában sajátos mezőgazdasági, erdészeti, egészségügyi stb. ágazati feladatokként jelennek meg, amelyek számítástechnikai megoldása során találkozunk az alkalmazandó matematikai algoritmusok és a biológiai tartalom egymáshoz rendelésével. A matematikai algoritmusok és a biológiai tartalom egymáshoz rendelésével és azok vizsgálatával foglalkozik a biomatematika.

A valószínűségszámításból kinőtt matematikai statisztika biológiai alkalmazása hozta létre a biometriát. A biometria hazai alkalmazásának egyik jelentős területe a mezőgazdaság volt, amelynek jelentősebb eredményeit dr. Hauser János foglalta össze, a ma már ritkaságszámba menő „Biometria” könyvében (Dr. Hauser János: *Biometria*; Debrecen, 1932).

A mezőgazdaságban elsősorban francia és angol alkalmazásokat vettek át és próbálták ki, illetve ezeket közölték. Mind nyilvánvalóbbá vált, hogy a biológiai feladatok maradéktalan megoldásához a biometria mellett szükség van a matematikának szinte minden ágára. A biológiai folyamatok matematikai modellezése és az alkalmazott matematikai fogalmak biológiai ellenőrzése hozta létre a biomatematikát. Addig, amíg a biometria hazai és nemzetközi szakirodalma igen nagy számú, a biomatematikáé, rövid múltjára tekintettel, kicsi. (Smith, C. A. B.: *Biomathematics, the principles of mathematics for students of biological science*, London, 1954; Kontar, B. A.: *A biomatematikai teóriák elemei és a technikai természetvédelem*. Ford. Moszkva—Szamarkand, 1968; Bán István: *Biomatematika és alkalmazása a növénytermesztésben*. Mezőgazdasági Kiadó., Budapest, 1979).

Az elmúlt 15 évben igen sok feladatot oldottunk meg, amelyek egy részénél hibákat is követtünk el. Szolgáljon ez mások okulására.

Mind a biomatematika, mind a biometria, alapjaiban matematikai módszereket és megfontolásokat tartalmaz. A gyakorlati alkalmazás sajnos sokszor túlzásokba esik, azaz „matematikaizmus”-ban szenved.

A biológiai jelenségek megismerésének egyik útja az, hogy a jelenségeket megközelítően azonos körülmények között figyeljük meg. A megfigyelések száma döntő tényezője a megbízhatóságnak. Sokszor a megbízható megismeréshez szükséges megfigyelések optimális száma nem kerül meghatározásra, és így fordulhat elő, hogy megdöbbentően kevés, vagy éppen feleslegesen sok megfigyelés történik. Adott esetben néhány megfigyelésből nagy horderejű következtetések születnek, illetve már az is előfordult, hogy tájékoztató jellegű ismeretek meghatározásához rendkívül nagyszámú megfigyelések történnek. Így fordulhatott elő, hogy egy tájékoztató jellegű, optimális mintavételi vizsgálathoz egy országos intézet minden megyei laborjában szinte a

teljes szagárda jelzett és nem jelzett kis magokat válogatott olyan mennyiségben, hogy egyes helyeken férőhelyül pajtát kellett kibérelni. Sajnos az is előfordult már, hogy kis számú megfigyelés után olyan hatóanyagok kerültek a mezőgazdasági és egészségügyi gyakorlatba, amelyek kifejezetten veszélyesek a környezetre és az emberre. Sajnos a káros hatóanyagok letiltását már jóval később, a szomorú és megdöbbentő tapasztalatok után rendelték el!

A legnehezebb feladat a matematikai és biológiai fogalmak helyes, kölcsönösen egyértelmű egymáshoz rendelése, így természetes, hogy a legtöbb félreértés itt tapasztalható. Egy részről biológiai jelenségekhez olyan matematikai fogalmakat rendelünk, illetve más részről matematikai paramétereknek olyan biológiai jelenségeket tulajdonítunk, amelyek a valóságban nem állnak fenn. Így gyakran előfordul, hogy a megfigyelt, de nem metrikus állapotjellemzőhöz (egyes tulajdonságok, formák, alakzatok stb.) önkényesen hozzárendelünk számértékeket, amelyek kiindulásul szolgálnak a matematikai feldolgozásnak. A növénytípológiában és a kórélettanban is előfordulnak olyan értékelési skálák, ahol az egyes, egymástól merőben különböző típusokhoz a skálának más-más értéke tartozik. A baj csak az, hogy a megfigyelések során feljegyzett skálaosztások segítségével akár egyszerű, akár súlyozott középértékeket számítanak. A kapott középértékeknek megfelelő skálaosztáshoz tartozó növénytípológiai vagy kórélettani típust fogadják aztán el átlagnak, amely pedig az adott környezeti lehetőségek között sokszor eleve kizárt. Az értékelési skálák kialakítására érdemes külön erre a célra kifejlesztett algoritmust használni. Az előzőeknek a fordítottja is előfordul, amikor is a matematikai paramétereknek téves biológiai jelenséget tulajdonítanak. Ha különböző hatóanyagokkal kezelt élő szervezet állapotjellemzőinek értékei alapján számított matematikai paraméterek megegyeznek, az nem jelenti például a különböző hatóanyagok hatómechanizmusának azonosságát.

A biológiai és matematikai megfeleltetések általában mindkét irányban együttesen szoktak előfordulni.

Gyakori eset a „kezelt” és „kezeletlen” halmazok közötti különbségek vizsgálata. A különbség kimutatására általában eleve csak a statisztikai próbákat (F , t stb.) használjuk fel, sokszor anélkül, hogy megvizsgálánánk a módszer alkalmazhatóságának feltételeit, pl.: normális eloszlás stb. Túl szigorúak vagyunk néhány feladatnál a szignifikanciaszinttel szemben, megkövetelnénk a $0,1$; 1^0_0 -os szignifikanciaszintet.

Biológiai vizsgálatok során — az ipari alkalmazásokkal szemben — adott esetekben a 10^3_0 -os szignifikanciaszintnek is örülhetünk. Vannak olyan esetek, amikor nem is kell erőltetni a különbségek kimutatására a statisztikai próbákat. Számtalanszor előfordul, hogy olyan égbekiáltó különbségek vannak, hogy kár bajlódni a statisztikai próbákkal. Gyakran előfordul az is, hogy igen kicsi — statisztikai próbákkal látványosan ki sem mutatható — különbségek vannak, viszont a különbségekben józan paraszti ésszel is felismerhető valamiféle tendencia. A kezelték például minden esetben vagy az esetek bizonyos százalékában kisebbek, mint a kezeletlenek. Ilyen, viszonylag gyenge állítások is igen hasznos információt jelentenek a további biológiai vizsgálatok számára. Vannak például olyan kórokozók, amelyek esetén kezelt és kezeletlen tenyészetek méretének vizsgálata során statisztikai próbákkal megbízható különbséget nem sikerült kimutatni, viszont felismerhető volt a kis különbségek tendenciájában néhány törvényszerűség. Ez alapján a biológiai vizsgálatokat nem vetettük el, hanem egyszerűen a tenyészeteknek más állapotjellemzőit mértük, pl. életképesség, szaporítóképletek jellemzői, újraoltások jellemzői stb., amelyek esetén már számítás nélkül is látszott a meggyőző különbség.

Ilyen és ezekhez hasonló esetek igazolják annak a kérdésnek a felvetését, hogy a biológiai vizsgálatok során milyen állapotjellemzőket figyelünk meg. Egyes esetekben a kísérleti gyakorlattal már eleve adottak a megfigyelendő állapotjellemzők. Más esetekben azonban meg kell határozni a megfigyelendő állapotjellemzőket. Először a biológiai kísérleti gyakorlatnak kell eldönteni, hogy mely állapotjellemzők meghatározására vannak meg a feltételek. Ezek közül kell kialakítani azokat, amelyek a megfigyelendő, ill. tanulmányozandó jelenséggel biológiai kapcsolatban levő állapotjellemzők körét tovább szűkíthetjük, más esetekben viszont éppen azt akarjuk felderíteni, hogy a kísérleti technika által megfigyelhető állapotjellemzők közül melyek vannak kapcsolatban az adott jelenséggel. Bonyolítja a kérdésfelvetést, hogy az egyes állapotjellemzők egymással is kapcsolatban vannak, és olyan, általunk nem ismert jelenséggel is, amiről nincs tudomásunk. Anélkül, hogy bonyolultabb, többszörösen összetett kapcsolatrelációkat említenénk, az előző biológiai vizsgálatok matematikai értékelései produkálják a legnagyobb meglepetéseket. Előfordult már, hogy a józan kísérletező észtt mellőzve, időt és pénzt nem kímélve, valamennyi, technikailag megfigyelhető állapotjellemzőt észleltük, előzetes matematikai vizsgálatok nélküli hatalmas mennyiségben. Ezek után a számítástechnikát kihasználva a többváltozós módszerek valamelyikével automatikusan lefuttatjuk az adott feladatot. Ilyenkor születnek a legmeglepőbb eredmények, a szomorú azonban az, hogy ezeknek hitelt is adunk és biológiai valóságnak fogadjuk el. Ilyenkor találkozunk komoly biológiai tudományos ülésszakokon olyan állításokkal pl., hogy egy adott növény terméseredménye nem függ a talajtól, vagy egy hatóanyag lebontása a szervezetben nem függ magától a szervezettől stb.

Az állapotjellemzők közötti kapcsolat leírására számos matematikai módszer áll rendelkezésre. Ezen módszerek általában a megfigyelt állapotjellemzők értékei vagy a megfigyelések alapján megállapítható matematikai tulajdonságok (különböző rendű diff.-hányadosok, integrálok stb.), illetve biológiai ismeretekhez rendelt matematikai fogalmak (számítási, mértani sorozat, rezgések stb.) alapján határozzák meg az állapotjellemzők közötti kapcsolat matematikai formáját.

A biológiai jelenség vizsgálatára használandó matematikai módszer kiválasztása sokszor eleve döntő hatással van a számítás eredményére. Sokszor a megoldandó biológiai feladat szempontjából olyan önkényes matematikai módszer alkalmazásával találkozunk, amely eleve téves információk megállapításához vezet. A leggyakrabban előforduló eset például a kétváltozós, regressziós függvény számításakor adódik, amikor is az előzetes biológiai összefüggés ismerete és a megfigyelési adathalmaz előzetes vizsgálata nélkül választják ki az alkalmazandó regressziós függvény típusát. Előfordul, hogy másodfokú görbét közelítő halmazt lineáris függvényvel, de az is, hogy másod-harmadfokú görbét közelítő halmazt negyed-, ötödfokú polinommal közelítenek. Ezen függvények illeszkedésvizsgálata, bár sokszor kielégítő, de mégis egész téves biológiai következtetések levonásához vezet.

A biológiában egyes megfigyelések esetén a biológiai állapotjellemző értékekre az értékelés során önkényesen matematikai analízisbeli függvényeket, ill. módszereket használunk. Egyes élettani EEG, EKG bioelektromos vizsgálatoknak számított matematikai analízisbeli függvények „kisimítják” a fűrészfogalmazokat, s ezáltal lényeges információk maradnak felderítetlenül.

A biológiai modellek és megfigyelések matematikai vizsgálata során kapott matematikai eredmények csak az adott biológiai környezetre, ill. biológiai feltételrendszerre érvényesek. Tévesek az olyan információk, amelyeket például

az adott biológiai vizsgálat értelmezési tartományára meghatározott matematikai modell alapján a megfigyelési értelmezési tartomány alsó korlátja alatti, illetve felső korlátja feletti tartományra állapítunk meg. Így egy kétváltozós élettani vizsgálat során, ahol a vizsgálat értelmezési tartományában a független változó növekedésével nő a függő változó, például az asszimiláció intenzitása, a hőmérséklet függvényében vizsgálva, T_1 és T_2 hőmérsékletkorlátok között monoton növekedő, akkor a megfigyelési értelmezési tartomány felső korlátja fölött már nem lehet állítani a monoton növekedést. A példában tehát T_2 hőmérséklet fölött nem biztos, hogy a hőmérséklet növekedésével nő az asszimiláció intenzitása. A vizsgálat értelmezési tartományán belül az alkalmazott matematikai módszertől függően eltérő biológiai következtetésekre juthatunk a kiszámított matematikai jellemzők alapján. Szélsőséges esetként fordulnak elő olyan példák, ahol egy szélső értékkel rendelkező halmazt szélső értékkel nem rendelkező matematikai modellel közelítünk. Ebben az esetben a matematikai modell paramétereit kiszámítva a biológiai következtetés is az, hogy a vizsgált értelmezési tartomány alsó vagy felső korlátjának közelében észlelhető a megfigyelési értékek szélső értéke, és hiába alkalmazunk egyébként szélső értékkel rendelkező matematikai modellt, a számított paraméterek és matematikai jellemzők ezt nem mutatják ki. Az előző példánál maradva, ha az értelmezési tartomány $(T_1; T_3)$, ahol $T_3 > T_2$ és a megfigyelési adathalmaz T_2 és T_3 között határozott maximumot mutat, akkor $(T_2; T_3) : (T_1; T_3) < \Sigma$ esetén a legkisebb négyzetek elvén kiszámított másodfokú parabola maximumhelye nem biztos, hogy a T_2 és T_3 közé fog esni. Ez biológiailag azt jelenti, hogy a hőmérsékletet T_3 -ig növelve, monoton nő az asszimiláció intenzitása. Előfordul azonban az is, hogy az előző példára harmadfokú polinomot alkalmazva, megtaláljuk T_2 és T_3 között a maximumhelyet, de T_1 és T_2 között egy minimumhelyet is találunk, mely biológiailag nem áll fenn. Hasonló helyzet áll fenn akkor is, amikor az alkalmazott matematikai modellek több megoldása lehet, pl. gyökhelyek, diff.-egyenlet megoldása stb., ilyenkor minden egyes matematikai megoldás nem jelent feltétlenül különböző biológiai jelenséget, hanem csak olyan lehetőségeket, amelyeket feltétlenül biológiailag ellenőrizni kell. Egy hatóanyag élettani hatását vizsgálva például, ha a megfigyelési halmazt negyedfokú polinommal lehet csak jól közelíteni (illeszkedésvizsgálat ekkor ad kielégítő eredményt) akkor a nullhelyek mindegyikének nincs biológiai jelentése. Egyes biológiai halmazok közelítésénél előfordul, hogy indokolatlanul magas fokszámú polinomot alkalmaznak, s bár a kiszámított polinom szinte minden megfigyelési értéken átmegy ($n-1$ -ed fokú polinommaximum), de biológiailag teljesen megalapozatlan, különböző lokális szélső értékeket vesz fel.

Nehéz feladatot jelentenek, és így gyakori félreértésre adnak okot a sokváltozós feladatok. A biológiai vizsgálatok nagy részében egyes megfigyelt állapotjellemzők igen sok más állapotjellemzőtől függenek, amelyek egymás hatását módosítják, valamint összetett hatások állnak fenn. Egyes matematikai modellek esetén (pl. egyes többváltozós regresszió modellekben) ha az egyes állapotjellemzők hatásának vizsgálata nélkül az összes, biológiailag hatástalan állapotjellemzőt is bevonjuk a matematikai modellezésbe, ez a feldolgozás kudarcát okozhatja (pl. fertőzések előrejelzése igen sok állapotjellemző alapja), s olyan biológiai következtetésre vezethet, hogy a vizsgált állapotjellemzők között nincs összefüggés. Ilyen feladatok vizsgálatára rendelkezésre állnak igen sok változós, speciális matematikai modellek, pl. többváltozós kiválasztási matematikai modell. Ritkábban fordul elő az az eset, amikor a vizsgált néhány biológiai állapotjellemző között határozott, jól modellezhető összefüggést ta-

lálunk, de annak oka egy általunk meg nem fejtett külső állapotjellemző, ill. jellemzők.

A hőmérséklet függvényében mért növényi bioelektromos potenciálváltozások esetében például a hőmérséklet-változás hatására megváltozott növényi anyagcsere-paraméterek megváltozása okozza a mért bioelektromos potenciál változását.

A biológiai jelenségek matematikai modellezését nagyban befolyásolja a számítástechnika. A számítástechnika előnyei azonban néha kevesebb gondolkodásra késztetik a vizsgálódó szakembert és az alkalmazható matematikai módszerek közül nem szakmai megfontolás alapján választjuk ki az adott feladat megoldására alkalmas egyetlen módszert, hanem a rendelkezésre álló módszerek mindegyikével lefuttatjuk a feladatot. A számítógépes futások eredményei alapján választjuk aztán ki a legjobb megoldást. Néha ilyen esettel találkozunk például a kétváltozós regressziós függvények kiszámításánál, amikor is egy futásban számítjuk ki a lineáris, másodfokú, harmadfokú, reciprokl, hatvány, exponenciális, logaritmus regressziós függvények paramétereit és egyéb matematikai jellemzőit.

Összegezve a leggyakrabban előforduló hiányosságokat, a biológiai jelenségek matematikai modellezésével soha nem szabad mellőzni a józan kísérletező ésszt és az állandó ellenőrzést.

A mobil aprítógépek műszaki fejlődése Észak-Amerikában — ellentétben Európával — kevésbé a tehergépkocsira szerelhető, 250—400 kW teljesítményű nagy aprítók irányába fejlődik, mint inkább a közép (forwarderekre épített, dobos és tárcsás) megoldások, valamint a kisebb berendezések, adapterek felé halad. Utóbbiak sorában jelentős figyelmet fordítanak a 35—90 kW tengelyvég-teljesítményű mezőgazdasági traktorokra három ponton felfüggeszthető aprítóberendezésekre. (FA., 80. 9. 186.)

Kéregből granulátum készíthető, amely úgy ég, mint a szén és azonos hőtermelés esetén csak fele annyiba kerül. A biomasszát aprítják, 20 százalékra szárítják és igen nagy nyomáson granulálják. (in SFw., 80. 273.)

Erdei utak tervezését számítógéppel kezdik végezni félautomata rendszer keretében a SzSzK-ban. A koncepció megoldása nyomán 1980-tól már javított programot adnak át szolgáltatás keretében a tervező erdőgazdaságoknak. Az eljárás a tervezés hatékonyságát és a megtakarításokat is növeli a beruházás és az építés terén egyaránt. (LPr., 80. 324.)

Kéregbrikett előállítására jelentett be szabadalmat egy osztrák mérnök. Az eljárás során nem használnak semmiféle kötőanyagot, a brikett csaknem teljesen elég, nem kormoz és 30 000 m³ kapacitású fűrészüzemben 1,5 millió liter fűtőolaj értékű ellentételt jelenthet. A prototípus építését most tervezik. A nagy értékű tüzelőanyagot bármilyen kályhában és központi kazánban, valamint távfűtő művekben bevezethetőnek tartják. (IHM., 80. 16/17.)

Szilfavész után tölgyfavész? Mind az erdővédelmi, mind a fafeldolgozó ipari szakembereket erősen foglalkoztatja nyugat-európaszerte az Észak-Amerikában már évtizedek óta elterjedő tölgyfavész behurcolásának kérdése. A főleg vöröstölgyeket károsító, de a mi őshonos tölgyeinken is fellépő, a szilfavész és a platánpusztulás kórokozójával közel rokon *Ceratocystis fagacearum* (Bretz) Hunt tömlösgomba fertőzése a korona felső részén jelentkezik a levélszélékből kiinduló barnulással. Az elszíneződés gyorsan halad és a vöröstölgyek nagyrészt még az évben elhalnak. A fertőzés rovarok és gyökérosszenővények útján terjed. Az Európai Gazdasági Közösség irányelvei a tölgy gömbfa Amerikából való behozatalának teljes letiltását írnyozzák most elő. Ez érzékenyen érinti a nyugateurópai fafeldolgozó ipart, mivel nagy mennyiségű import tölgyvel dolgozik. Az amerikaiak túlzottnak tartják az aggodalmat, miután náluk főként vöröstölgyet károsít a vész, az európaiak pedig érthetően védeni kívánják saját faállományait. Különös jelentőségű ez Franciaország tekintetében, itt az erdők 34%-a tölgyes. A kérdést a vész biológiájának részletesebb vizsgálatával kívánják most eldönteni. (FORSTARCHIV 1979. 7/8. Ref.: Jérôme R.)