Szám és kép
Grafika az adatfeldolgozásban

Szepesváry Pál*

A számokat valószínűleg azóta kísérik képek, amióta a kultúrában felbukkant a számok tana. Egyiptomi papíruszon csakúgy találunk matematikai gondolatmeneteket szemléltető ábrákat, háromszögeket, köröket, mint a babilóniai agyagcserepeken [1]. De, nagyot ugorva az időben, Descartes is képekkel szemléltette az összeadás, kivonás, szorzás sőt az osztás lényegéről szóló értekezését [2], oly módon, hogy azt, háromszáz évvel később, korszerű könyvek [3] is változatlanul használhatták. Galilei a „Matematikai érvelések..."-ben [4], már annyi képpel szemléltet, mint egy ma megszokott műszaki kézikönyv szerzője.

Nyilvánvalóan nem véletlen, hogy egy műszaki ember ma matematikára támaszkodó szöveget képek nélkül elképzelni se tud. (Nem tartozik ugyan tárgyunkhoz, de elgondolkozhatunk azon, mennyire nincs ez így az irodalmi, filozófiai, politikai szövegekben.) A „literary intellectual" H.G. Wells, ismervén a tudomány embereinek eszejárását is, kedves iróniával éppen azzal jellemzi „minden nőni tudót megmérő" hősét, Redwoodot, hogy az egyenletes növekedést „saját kifejezésmódja szerint" ——————— vonallal, az ugrásokkal, megszakításokkal járót pedig „valahogy így":

jellemzi [5].

A rajznak, képnek mindennapos használata szinte természetesen kételyeket is fölvet. Abban senki sem kételkedik, hogy egy gépalkatrész, vagy egy épület műszaki rajzának mondanivalóját szövegesen nemhogy röviden, hanem – az emberi nyelv korlátai miatt – egyértelműen egyáltalán nem sikerülne visszaadni. Ugyanakkor gyanakodnunk kell arra, hogy az ábrákkal, képekkel közvetített szemlélet be is csaphat. Amennyire elterjedt a szemléltetés mindennapjainkban, annyira óvnak a matematikusok a szemlélet szeretetétől. Ezt az ellentmondásos helyzetet érdemes elemeznünk.

Ismerjük fel, hogy munkánk során rajzot, képet két módon használhatunk: *számításra* és *szemléltetésre*. Mindkét esetben rosszul is járhatunk, de más-más módon. Ha tisztázzuk szándékainkat, számos vitának elejét vehetjük.

Számítások rajzokkal

Az iparban, mezőgazdaságban ma alkotó szakemberek alighanem már ritkán találkoznak azokkal a vaskos grafikongyűjteményekkel, amelyek két generációval ezelőtt mindennapos munkaeszközök voltak az íróasztalokon. Legyen szó kőolajtermékek sajátságairól különböző feltételek között, vagy szerkezeti anyagok tulajdonságairól, netán statikai feladatokról, a számolóábrák kezünk ügyében voltak. Az ábra, a vonalzó, esetleg valamilyen leleményes leolvasó eszköz hamar eljuttatott a várt eredményhez.

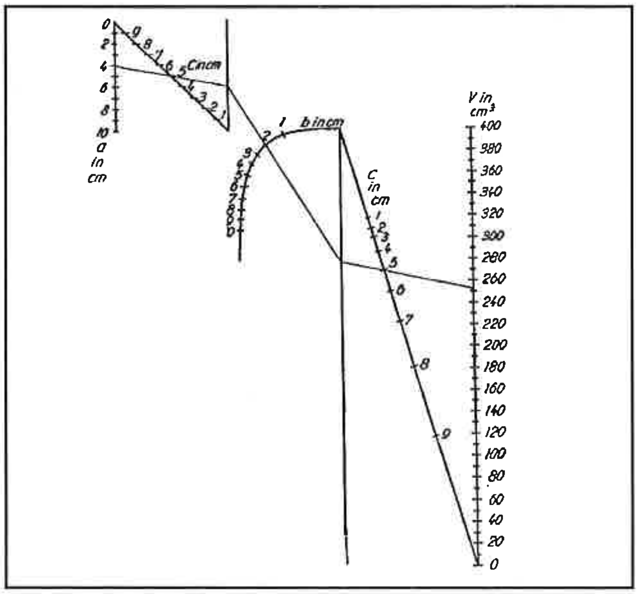
Az ábrák készítése, a grafikus számítástechnika szakmává, tantárggyá, tanulmányok témájává vált. Akkor talán nem volt annyira világos, mint ma, hogy ez a szakma, kissé gonoszul szólva „pótcselekvés", abban az értelemben, hogy azért kellett, mert hiányoztak a gyors módszerek, magyarán szólva a számítógépek, a többváltozós nemlineáris feladatok ma is csak numerikusan megoldható feladatainak megoldásához. Mindez persze nem von le semmit annak értékéből, hogy ezekre a módszerekre szükség volt és megalkotóik tiszteletünket érdemlik.

A számolóábrák a függvények képeiből fejlődtek ki. Egyszerű grafikonokról olvastunk le érdekes adatokat, mondjuk egy S-alakú görbe inflexiós pontjának abszcisszáját. Ha a kép görbeseregekből állt, a használhatóság nyilvánvalóan nőtt. Az ábrák megrajzolását, de a leolvasást is könnyítette, ha az egyenletesen beosztott tengelyeket esetenként másfajta, sűrített, ritkított, logaritmizált stb., beosztással cseréltük fel. Úgy, ahogyan ez a mai napig is szokásos pl. bizonyos S-alakú görbéknek logit-transzformációval való kiegyenesítésénél [6].

A számolóábrák leglelemenyesebb változatai voltak a *nomogramok*, amelyek készítése kemény erőfeszítésekbe került. Nem szándékunk, talán értelme sem volna, hogy a nomográfia részleteinek ecsetelésébe bocsátkozzunk. Azoknak, akik nomogramokkal nem találkoztak, annyit, hogy a nomogramban a változókat legkülönbözőbb alakú és beosztású vonalak képviselték, amelyek pontjait egyenesekkel összekötve, adott metszéspontokon leolvashatók voltak a keresett eredmények. Az *1. ábrán* (példaképpen) bemutatott nomogram [7] két ellipszoid térfogatának kiszámítására való, amelyek közül az első főtengelyeinek hossza *a*, *b* és *c*, a második (forgási) ellipszoidé *b* ill. *10 – c*. A szóban forgó

$$V=\frac{4}{3}\pi\left[abc+b^2(10-c)\right]$$

összefüggés baloldala kiszámításának programja ma személyi számítógépen egyetlen sor, egy tenyérbe simuló, programozható zsebszámológépen 15-20 utasítás, a kiszámítási idő pedig polgári időmérő eszközökkel nem is mérhető.



1. ábra. Nomogram ellipszoidok térfogatának kiszámítására

A számolóábrák taxonómiájának tárgyalása során is eljuthatunk a mozgó skálájú eszközökhöz: a logarléchez, a logaritmikus tárcsákhoz. Ezek felett – tagadhatatlanul bizonyos nosztalgiával kísérve – csakúgy eljárt az idő, mint a számolóábrák felett. Számolóábrákra szükség volt, de nehéz volt a szerkesztésük, nem volt kiváltképpen egyszerű a kezelésük, és végül is korlátos a pontosságuk.

Szemléltetés rajzokkal

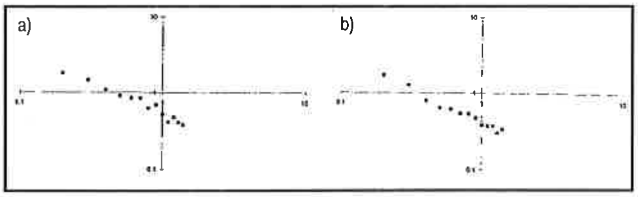
Éppen a bevezetésben említett történelmi tapasztalat utal arra, hogy valahol az ember biológiai, érzékelésbeli adottságaiból, a szem páratlan alakfelismerő képességéből következik, hogy matematika rajzos szemléltetése olyan eredményes.

Ahogyan az ergonómusnak tudatosan segíteni kell az eligazodást piktogramokkal, ugyanúgy érdemes grafikusan felhívni a figyelmet olyan jelenségekre, amelyek számokból, számtáblázatokból nem tűnnének ki feltűnően. A lehetőségek választéka bőséges, már-már óvatosságra van szükség.

Egy függvény grafikonjából már a kezdő is látja az ábrázolt folyamat első néhány deriváltját: emelkedő, állandó, süllyedő jellegét, sebességét, gyorsulását, lassulását, sőt némi gyakorlat után már megsejti, mely formula írja le a szemlélt összefüggést. Bonyolódik azonban a helyzet, ha a grafikon tengelyeit transzformálják és különösen kritikussá válhat a szemléltetés haszna, ha valaki a rajzot saját, alkalmasint alaptalan várakozása igazolására akarja felhasználni.

Tudott például, hogy a logaritmikus transzformáció mennyiséggel arányosan csökkenti a mért értékekhez kapcsolódó bizonytalanságot, nagyobb súlyt kölcsönözve ezáltal a nagyobb értékeknek, ugyanakkor egy kisebb érték kisebb hibája nagyobb súllyal esik latba. Bizonytalan értékek alapján megrajzolt grafikon becsaphatja a szemlélőt. A *2. ábrán* bemutatott, kétszer logaritmikus léptékben ábrázolt két pontsort aligha lehet megkülönböztetni. Mindkét esetben hajlanánk arra, hogy a pontok közé egyenest húzunk, holott az első (a) pontsor egy $a{\cdot}x^{b}$ függvényhez, a második (b) egy ae^{-bx} függvényhez tartozik. A pontok $\pm$15 relatív százalékos, egyenletes eloszlású véletlen hibákkal vannak szórva. A szemlélettel szemben helyes a kötelesség kétkedést tanúsítani, a függvények vélt

* ELTE Kémiai Tanszékcsoport, Budapest.

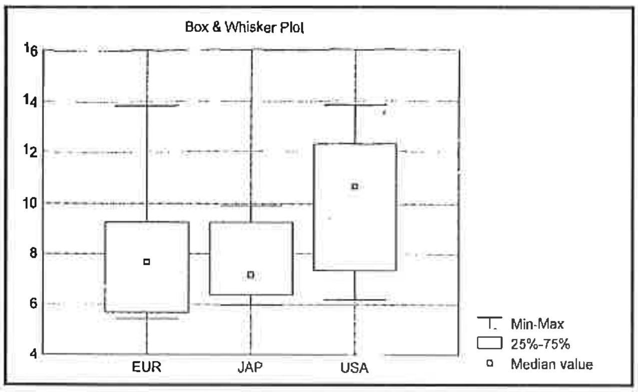


2. ábra. Hatványfüggvény és exponenciális függvény megszórt pontjai logaritmizált tengelyek között

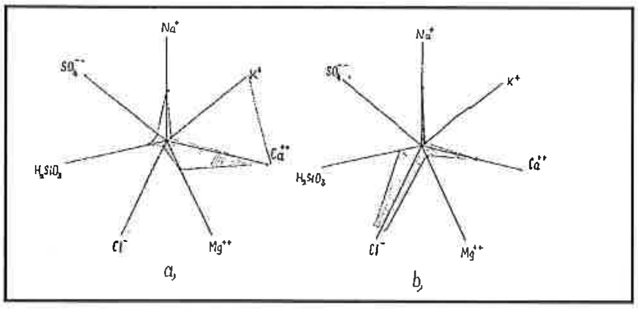
alakjaira vonatkozó feltevésekről végezzük el a bőségesen kínált, reziduálisokra vonatkozó szignifikancia vizsgálatokat. [7–9].

Különösen hasznos és hatásos a képpel való szemléltetés akkor, amikor olyan objektumokat (létező tárgyakat, mérési eredményeket stb.) kívánunk szemügyre venni, amelyek jellemzői szórnak, más szóval, amikor statisztikai sokaságokat tanulmányozunk. Tudott, hogy bizonyos, jól definiált statisztikai leíró mennyiségek, mint pl. átlagok, mediánok, szórások, terjedelmek valók ara, hogy fogalmunk legyen a tekintett adathalmazról. Tudott az is, hogy az adathalmaz elemek adott mérethatárok közé eső hányadainak oszlopos ábrája, a hisztogram már szemléletes kép az adatok eloszlásának jellegéről, szimmetriájáról, kiterjedéséről. Kevésbé ismert, holott különösen hasznos ábrázolás e szóban forgó célra a *szakállas ábra*, a *box and whiskers* diagram.

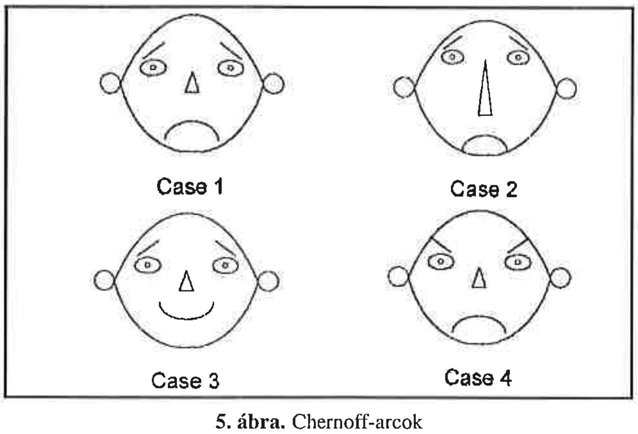
Mielőtt az ábráról részletesebben beszélnénk, emlékeztetünk arra, hogy egy adathalmaz jellemző adatának tekinthető a *minimuma*, azaz legkisebb elemének értéke, és a *maximuma*, továbbá azok a határok,



3. ábra. Európai, japán és amerikai kocsik üzemanyag fogyasztása



4. ábra. Két ásványvíz összetétele



5. ábra. Chernoff-arcok

amelyeknél az elemek tekintetbe vett értéke az elemek 25, 50 illetve 75 százalékában kisebb (*első kvartilis, medián, harmadik kvartilis)*.

Ezeket az értékeket szemlélteti az említett boksz and whiskers ábra, amely nevéhez híven egy „dobozból" és a belőle kinyúló „pofaszakállakból" áll. A doboz téglalap, amelynek magassága a harmadik és első kvartilis közötti interkvartilis távolság, közepén bejelöltetik a medián, a szakállak pedig a minimumtól a maximumig nyúlnak. Nyilvánvaló, hogy az ábra a számértékeken túl az eloszlás esetleges aszimmetriáját is jól szemlélteti, és jól használható különböző sokaságok összehasonlítására. A *3. ábra* európai, japán és egyesült államokbeli személygépkocsik 100 km távolságon fogyasztott üzemanyag mennyiségét hasonlítja össze.

A szakállas ábrák léptékének a bemutatott csak legegyszerűbb változata. Van például olyan konvenció is, amely szerint a szakállak ne a minimumig és a maximumig nyúljanak, hanem csak a doboz fölé és alá felmért interkvartilis távolságon belül fekvő legnagyobb és legkisebb értékig. Az ezeknél nagyobb és kisebb értékek ábrázoltassanak *-gal és tekintessenek kilógó értékeknek, outlier-eknek [10].

Hasonlóan segítik különböző adatsorok gyors összehasonlítását a csillagábrák. Az adathalmazokkal leírt objektumok tulajdonságainak feleltessünk meg origóból sugarasan kiinduló tengelyeket, amelyekre fölmérjük a szóbanforgó tulajdonság mérőszámát. A *4. ábrán* két magyarországi ásványvíz (a: székesfehérvári, b: miskolc-tapolcai) fontosabb összetevőinek koncentrációi szerepelnek, mint pontok, a sugarakon, amely pontokat azután egymással összekötöttük. Belátható, hogy gyakorlott szemű elemző kémikus az ilyen ábrákon kialakuló csillagok körvonalai alapján rápillantással osztályozni képes ásványvizeket.

Adathalmazok jellemző tulajdonságainak legszellemesebb és szinte leghatásosabb képi szemléltetését azonban alighanem a *Chernoff-arcok* jelentik. Onnan kell kezdenünk, hogy az emberi arc végtelen változatosságban képes tükrözni hangulatokat, indulatokat, kedvet, szomorúságot és ezt a metakommunikációt mindenki érti. Egy asztal mellé ülő ínyenc arcán éppen úgy tükröződik az elé helyezett tányérban kínált étek összetétele, mint ahogyan egy üzemirányító arcán gond vagy nyugalom tükröződik, ha vizsgálja üzemének mérési adatsorát. Ezt modellezi, szimulálja az a program, amely a Chernoff-arcokat szerkeszti meg. Ha tudjuk, hogy valamely adat számunkra kedvezőtlen, rajzoltassunk olyan arcot, amelyen az ennek az adatnak számértékével arányosan magasra szökik a szemöldök. Ha egy másik adat derűre ad okot, húzódjon fülig a Chernoff-arc szája. Az *5. ábrán* játékos bemutatásképpen egy édességkedvelő antialkoholista arca szerepel, akinek első esetben átlagos, második esetben húsos, harmadik esetben édes, negyedik esetben pedig italos menüt kínálnak.

A Chernoff-arcok rendkívül árnyaltan programozhatók. Az adathalmazban szereplő számértékektől függően, adott lépésközökkel beállítható az arc szélessége, szimmetriája, a fülek nagysága, helyzete, az orr hossza, szélessége, a száj helyzete, hossza, görbülete, a szem magassága, távolsága, szimmetriája, ferdesége, a pupillák helye, a szemöldök magassága hossza, szöge... összesen húsz jelleg. Nem nehéz elképzelni, hogy kellő türelemmel és beleéléssel programozva és kellő tapasztalatot megszerezve valamely adatrendszer mondanivalóját a pillanat törtrésze alatt fel tudjuk fogni.

A tanulság

Matematikusok, adatfeldolgozók, műszakiak, felhasználók találkozásakor, bármily különös, fölparázslik a vita, kell-e, szabad-e következtetéseink során képekre támaszkodnunk. Az elmondottak alapján azt gondolom, láthatjuk a vita nem alaptalan, a képek segíthetnek, becsaphatnak. Mint mindenhol, ész nélkül eljárni súlyos hiba. Ugyanakkor, mint mondtuk, éppen az emberi érzékelésbeli adottságok folytán, hallatlan segítség egy sikerült ábra. Olyan segítség, amely nincs is távol. Mindenki számára hozzáférhető programok tartalmazzák a rajzoló eljárásokat, az egyszerű függvényábrákon kezdve a Chernoff-arcokig. A lehetőségek asztalunkon adottak [11].

Források:

[1] *Simonyi Károly:* A fizika kultúrtörténete. Gondolat Kiadó, Budapest, 1978. pp. 35-40.
[2] *Descartes:* Válogatott filozófiai művek. I. Szabályok az értelem vezetésére, XVIII. Szabály. Akadémiai Kiadó, Budapest, 1961. pp. 148-153.
[3] *Korwien, H.:* Graphisches Rechnen Nomographie. Fachbuchverlag GmbH, Leipzig, 1949. pp. 9-11.
[4] *Galiei, G.:* Matematikai érvelések és bizonyítások. Európa Könyvkiadó, Budapest, 1986.
[5] *Wells, H.G.:* Az istenek eledele, Szépirodalmi Könyvkiadó, Budapest, 1954. p. 12.
[6] *Frank I. és Todeschini, R.:* The Data Analysis Handbook, Elsevier, Amsterdam, 1994. pp. 334.
[7] *Kemény S. és Deák A.:* Kísérletek tervezése és értékelése. Phare, Budapest, 1998.
[8] *Sachs, L.:* Angewandte Statistik. 7. Auflage. Springer, Berlin, Heidelberg, 1992. p. 373-420.
[9] *Vincze, I.:* Matematikai statisztika ipari alkalmazásokkal. 2. kiadás. Műszaki Könyvkiadó, Budapest, 1975. pp. 110-163.
[10] *Sachs, L.:* Statistische Methoden. Planung und Auswertung. 7. Auflage. Springer, Berlin, Heidelberg, 1993. pp. 39-40.
[11] STATISTICA for Windows. StatSoft, Inc. Release 4.5, 1993.