II.F1. függelék
Alapismeretek az AS-EASY-AS... táblázatkezelő program 5.50c verziójának használatához

A táblázatkezelő programok a legelterjedtebb számítógépes alkalmazások a szövegszerkesztők után. Ennek alapvető oka hatékonyságuk. Segítségükkel rengeteg olyan feladat megoldható, ami számításokat igényel, így természettudományos problémák megoldásában is fontos eszköz lehet.

Az ilyen típusú programok használatánál a legalapvetőbb fogalom a munkalap. Ez semmi más, mint egy kitöltésre váró táblázat, de a munkalapot akár egy mátrixként is felfoghatjuk, amelynek minden elemét külön-külön tudjuk kezelni. A munkalap alapegysége a cella, ez megfelel egy táblázat adott oszlopában és sorában lévő rovatnak (vagy egy mátrix egy elemének). A hasonlóság a táblázatokkal abban is fennáll, hogy a cellára is úgy hivatkozunk, hogy megadjuk az oszlop- és sorszámukat. Az oszlopot egy vagy több betűvel jelöljük (A, B, ..., Z, AA, AB, ..., AZ, BA, stb.), míg a sorszám valóban egy szám (1, 2, 3, ...). Így az F37 jelölés a munkalap 6. oszlopában és 37. sorában található cellát jelöli.

Egy cellába egy adatot tudunk beírni, ami lehet számadat, karaktersorozat, vagy kifejezés. Ez a kifejezés sokminden lehet, pl. egy matematikai függvény definíciója is. A definícióban hivatkozhatunk más cellák tartalmára és ez a tulajdonság teszi hatékonnyá a táblázatkezelő programokat. A cellákban tehát nemcsak egyszerű adatokat tudunk tárolni, hanem adatok közötti összefüggéseket is. Ha egy számot tartalmazó cellában megváltoztatjuk a számértéket, akkor az összes olyan cella tartalma is megváltozik, amely hivatkozik a megváltoztatott értékű cellára. Ha pl. egy függvény értékét több pontban akarjuk kiszámítani, akkor a műveleti jeleket minden egyes alkalommal meg kell ismételnünk, ha számológépet használunk. Táblázatkezelő program alkalmazása esetén erre nincs szükség, a függvény definícióját csak egyszer kell beírni.

Ez az elv Bostonban született 1980 körül. A világ első táblázatkezelő programját — VisiCalc néven, MacIntosh számítógépre — hobbiból írta egy egyetemi tanár. Eljárását nem szabadalmaztatta, így gombamódra szaporodtak a különböző cégek által írt programok, amelyek közül kettő emelkedik ki:

– A **Lotus 123** az egyik legjobban sikerült program. Sokáig szinte kizárólagos etalon volt, a legtöbb cég terméke (legalábbis külsőleg) erre próbált hasonlítani. Mai változatai is a legtöbbet tudó és a legkevesebb rejtett hibát tartalmazó programok a piacon.
– A **Microsoft Excel** sokak szerint a Microsoft cég legjobban sikerült programja. Főként azzal hódította meg a piac nagy részét, hogy a Lotus 123-nál jóval könnyebben kezelhető és az adatcsere más programokkal biztosabbnak ígérkezik.

A laboratóriumi gyakorlatokon mért adatok kiértékelésében is sok segítséget nyújthat egy táblázatkezelő program. A kiértékeléshez a Lotus 123 egyik jól sikerült utánzatát, az **AS-EASY-AS...** nevű táblázatkezelőt biztosítjuk, mert ez a program a legszerényebb képességű IBM PC kompatibilis számítógépen is megbízhatóan fut. A függelék a táblázatkezelő program használatához szükséges alapvető ismereteket foglalja össze.

Az **AS-EASY-AS...** 5.50c verziója egy DOS operációs rendszer alatt futó táblázatkezelő program. A programhasználat és képernyőfelépítés megfelel a Lotus 123-, illetve mindazon programokénak, amelyek annak megfelelően képesek működni (pl. Quattro Pro, Excel stb.). A munkalap maximálisan 255 oszlopból és 8192 sorból állhat. A munkalapokat a program olyan formában menti el (**.WKS** kiterjesztéssel), hogy azok minden további változtatás nélkül használhatók minden olyan táblázatkezelővel, amely érti a Lotus 123 formátumát.

Az **AS-EASY-AS...** indítása

A hallgatók számára beállított gépeken a program az **EASY** parancs begépelésével indítható. Az **ASEASY.EXE** program közvetlen indítása nem javasolt, mert az indítást végző — általunk írt — **EASY.BAT** program olyan paramétereket is tartalmaz, amelyek nincsenek benne a program konfigurációs adatállományában. A szükséges paraméterek hiányában működési rendellenességek léphetnek fel bizonyos feladatok elvégzése során. Az **EASY.BAT** azt is biztosítja, hogy a billentyűzet ugyanolyan módon legyen használható, mint a program korábban legelterjedtebb, 4.0-ás verziójában. Ha valaki önmagában indítja a programot, az legyen tisztában a használható DOS-paraméterekkel.

A képernyő értelmezése

Az **ASEASY.EXE** betöltés után megjeleníti a címoldalt, majd megáll. Bármely billentyű lenyomása eltünteti a címoldalt, és egy üres munkalapot nyit meg. A képernyő legnagyobb részét a még üres cellák (más néven mezők) foglalják el. Ezek közül az aktív mező eltérő színű (fekete-fehér képernyőn világos), gépeléskor ebbe a mezőbe tudunk adatokat írni. A képernyőlap egyes részeinek a következő a funkciója:

– A legfelső sor elején látható az aktív mező neve, azaz elöl az egy vagy két betűs oszlopjelölés, utána pedig a sorszám. A1 például azt jelenti, hogy a munkalap bal felső sarkában lévő mező az aktív. A nevet egy kettőspont követi, amely után a mező aktuális tartalma látható.

  A képernyő legfelső sora egyben parancssor is. Ha az aktuális mezőbe valamit gépelni kezdünk, az ott jelenik meg. Oda kerül a mező tartalma akkor is, ha javítani kívánjuk. Ez az F2 (Edit) billentyűvel érhető el.

- A képernyőlap második sorában láthatók az oszlopokat jelölő betűk, a bal szélen pedig a sorok sorszámai. Az aktuális mezőnek megfelelő oszlop- és sorjelölések eltérő vagy inverz színben jelennek meg.
- Alulról a harmadik sor mutatja a munkalap számát (egyszerre több munkalap is lehet a memóriában), és a munkalapot tartalmazó adatállomány nevét is (ha már egyszer elmentettük). Ebben a sorban és a képernyő jobb oldalán találhatók azok a területek, amelyek az egér segítségével mozgatják a munkalap képernyőn megjelenő részét.
- A képernyőlap alján az utolsó előtti sorban az egyes funkciógombok (F1-F10) feladatának rövidített megjelölése látható.
- A képernyő legalsó sorában a program és a számítógép aktuális állapotáról kapunk információkat:
  a) A program verziószáma után (V55) a rendelkezésre álló szabad memóriát látjuk százalékban és kilobyte-ban is megadva.
  b) Ezután egy **`Auto`** vagy **`Man`** felirat jelenik meg. Az előbbi esetén minden cella tartalma automatikusan újraszámolódik, ha egy cella tartalmát megváltoztatjuk. Az utóbbi esetben ez nem történik meg, az újraszámolás az F9 (Calc) funkcióbillentyűvel érhető el.
  c) Az alsó sor közepén láthatjuk a program aktuális állapotát. Itt sokféle felirat lehet, ami arról tájékoztat, hogy mit csinál éppen a program. Itt nem soroljuk fel az összes üzenet csak a legfontosabbakat. A **`READY!`** felirat azt jelenti, hogy a program éppen nem csinál semmit, csak várakozik a felhasználó parancsaira. A **`+VALUE`** azt jelzi, hogy éppen egy számadatot gépelünk be, míg a **`` `LABEL ``** üzenet akkor jelenik meg, ha egy karaktersorozatot írunk egy cellába. Makroprogramok futtatása alatt a **`MACRO !`** feliratot láthatjuk. Ha éppen a menürendszert használjuk (ld. alább), akkor itt jelenik meg az is, hogy a menürendszer melyik ágában vagyunk (pl. "**`-File-`**" üzenet).
  d) Az alsó sor további részében az egyes speciális billentyűk állapotáról kapunk információt, nevezetesen arról hogy *(1)* az End billentyűt nyomtuk-e meg előzőleg (ld. alább), *(2)* átváltottunk-e felülírás módra az Insert gombbal (ekkor az **`Ovr`** üzenet látható) és érvényben van-e a *(3)* Num Lock, *(4)* Caps Lock vagy a *(5)* **`Scroll Lock`** (jelentése az **AS-EASY-AS...**-ben: irányított kurzormozgatás) beállítás.
  e) Végül a státuszsor jobb végén a számítógépen beállított időt mutatja egy óra.

Mezők kiválasztása és adatok bevitele

A mezők között a kurzormozgató billentyűkkel (nyilakkal) tudunk mozogni. Hatásosak a PageUp és PageDown billentyűk is, amelyek 20–20 sornyit mozdítják el a kurzort. Oldalirányban jobbra a Tab, balra a Shift-Tab kombinációval léphetünk egy képernyőnyit. A Home billentyű lenyomásával visszajutunk az A1 mezőbe. Az End billentyű lenyomása módosítja a közvetlenül utána lenyomott kurzormozgató billentyű működését (ehhez az End-et nem kell végig lenyomva tartani!). Az End lenyomása után bármelyik kurzormozgató billentyű hatására a megfelelő irányban a blokk szélére kerül a kurzor. A program a blokk kifejezés alatt a kitöltött mezőket érti, azaz az adott irányba eső első kitöltött mezőt, ha az aktuális mező üres; illetve az utolsót, ha az aktuális mező kitöltött. Az F5 funkcióbillentyű lenyomásával aktivizálható a GOTO parancs, amely számára bármely mező nevét megadva a kurzor egyenesen arra a mezőre kerül.

A mezők kitöltése nagyon egyszerű. Egy mezőre ráállva abba közvetlenül számot, szöveget vagy matematikai összefüggést írhatunk be. Az Enter/Return vagy bármely kurzormozgató billentyűvel jelezhetjük, ha befejeztük a gépelést. A beírtakat a program elsősorban számként igyekszik értelmezni. Ez sikerül, ha a beírt karakterek csak számok és műveleti jelek(pl. **`12*5+7/2-50*1.5E-1`**). Ha ez nem igaz, akkor a program automatikusan balra igazított szövegként kezeli a beírtakat, és az általunk beírtak elé egy aposztróf (`'`) karaktert ír. A begépelt karakterek akkor is szövegként íródnak be, ha csak számokból és műveleti jelekből állnak, de az első beírt karakter egy aposztróf. Az aposztróf helyett idézőjelet (`"`) írva jobbra igazított szöveget kapunk, míg a középre igazítás a hatványjellel (`^`) történő kezdéssel lehetséges.

Matematikai összefüggésként igyekszik kezelni a program minden olyan karaktersort, amelyik a **`+`**, **`-`**, **`(`** vagy `@` karakterek valamelyikével kezdődik. Az ilyen sorok szintaxisát akkor ellenőrzi a program, amikor befejeztük a beírást. Ha a szintaxis megfelel egy matematikai összefüggésnek, akkor ez bekerül a mezőbe, egyébként a program hibajelzést ad és nem enged továbblépni. A **`+`** és **`-`** értelemszerűen előjelet jelent, míg a `@` karakter azt jelzi, hogy valamelyik beépített függvény neve következik. Ennek argumentuma(i) közvetlenül a név után zárójelek között található(k) (pl. **`@SIN(3.14)`** vagy **`@SUM(E2..E14)`**).

Sokféle előre definiált függvény érhető el. A matematikai, statisztikai, logikai és közgazdasági függvények mellett a szövegkonstansokkal végezhető eljárások is megtalálhatók az F1 (Help) billentyűvel előhívott Súgóban. A képletekben számok, függvények és műveleti jelek mellett mezők nevei is szerepelhetnek. Ekkor az adott mező értékére vonatkozik az elvégzendő feladat, attól függetlenül, hogy az egy eredetileg számként beírt adat, vagy már egy másik számolás eredményeként keletkezett.

Bonyolult összefüggéseket kétféle eljárással tudunk az adott mezőbe beírni:

1. Az egyik a közvetlen begépelési mód, amikor a mezők nevét is a billentyűzetről visszük be. Ehhez előre tudnunk kell a mezők nevét.
2. A másik eljárás nem teszi szükségessé ezt az előrelátást. Az összefüggés beírása során megmutathatjuk a programnak, hogy mely mezők között végezze el a műveletet. Ezek neve automatikusan beírásra kerül a matematikai összefüggésbe.

Például, ha egy mezőbe azt akarjuk írni, hogy az tartalmazza a B3 és a C1 mező összegét, akkor az első megoldás szerint a mezőre ráállva begépeljük a **`+B3+C1`** karaktersorozatot. A második eljárásnál először a mezőre állunk, majd egy **+** begépelésével jelezzük, hogy matematikai összefüggés következik. Ezután a kurzormozgató billentyűkkel elmegyünk a B3 mezőre. Ekkor a képernyő legfelső sorában a **`+B3`** beírás szerepel. Újabb **+** jel begépelése hatására a kurzor visszaugrik az eredeti célmezőre és a felső sorban a **`+B3+`** kifejezés jelenik meg. Ekkor a kurzormozgató billentyűk segítségével elmegyünk a C1 mezőre így a **`+B3+C1`** kifejezés lesz látható a parancssorban.

Mivel a beírandó kifejezés készen van, az Enter/Return billentyűvel jóváhagyjuk. Ha szintaktikusan helyes a kifejezés, akkor a kurzor visszaugrik az eredeti célmezőre és megmutatja a művelet eredményét. *Figyelem! Az Enter/Return billentyű helyett a kurzormozgatók nem használhatók matematikai kifejezés beírásakor*, mivel azok az elmozdulásnak megfelelően megváltoztatják az utolsó mező nevét, így nem zárhatják le a beírást.

Mezők tartalmának másolása, mozgatása

Nem sok segítséget jelentene a táblázatkezelő, ha minden elvégzendő, egymással analóg számítást egyenként kellene elvégeztetni. Ilyen számítás lehet például, ha két vagy több oszlopban lévő adat között soronként ugyanazt a műveletet kell elvégezni.

Táblázatkezelő program alkalmazásával a fenti típusú probléma megoldásához csak egy célmező esetében szükséges az összefüggés definiálása, a többit rábízhatjuk a program CopyCell parancsára.

A CopyCell parancs a menüből adható ki. A menü kinyitására a törtjel (/) billentyű szolgál. Álljunk rá a másolni kívánt mezőre, majd nyomjuk meg a törtjelet. A menüből válasszuk ki a kurzormozgatókkal a CopyCell sorát és hagyjuk jóvá az Enter/Return billentyűvel (ezek helyett egyszerűbben a C billentyűt is nyomhatjuk a törtjel után, ami a CopyCell parancs ún. gyorsbillentyűje). A parancssorban megjelenik annak a mezőnek a neve, amelyiken állunk. Ha a kurzort elmozdítjuk, akkor annak a másik mezőnek a neve is megjelenik két ponttal elválasztva, amelyre elmozdultunk (pl. **`A3..E9`**). Tetszőleges téglalap alakú területben lévő mezőket másolhatunk. A fentiekből is látszik, hogy egy tartományra (a jelenlegi esetben a másolandóra) úgy hivatkozhatunk, hogy két ellentétes "sarkában" lévő mezők nevét két ponttal összekötjük. Ha kijelöltük a másolandó területet, akkor az Enter/Return billentyűvel jóváhagyjuk.

Ezek után a parancssorban megjelenik egy felirat, ami azt mondja meg, hogy hová akarjuk másolni a kijelölt mezőket. Itt kezdetben mindig az aktuális mező neve jelenik meg. A kurzorral elmozdulva menjünk az első olyan mezőre, ahová másolni kívánjuk a korábban kijelölt részt. Ha csak egyetlen másolást kívánunk elvégezni, akkor hagyjuk jóvá a választást az Enter/Return billentyűvel. Ha azonban többször, azaz egy egész tartományba kívánunk másolni, akkor nyomjuk meg a pont (.) billentyűt. Ezzel "lehorgonyozzuk" a kijelölendő terület egyik sarkát. A kurzorral mozogjunk a célterület másik, átlós sarkába. A parancssorban ez is két ponttal elválasztva jelenik meg az előző sarok címétől. Jóváhagyva a másolás megtörténik.

Egy egyszerű példa a fentiekre. Ha pl. a **+A1+B1** kifejezést, amely a C1 mezőben volt, átmásoltuk a C2-től D5-ig lévő tartományba (rövidebben C2..D5), akkor C2-ben a **+A2+B2**, C3-ban a **+A3+B3**, stb. kifejezések lesznek:

	A	B	C	D
1	?	?	+A1+B1	
2				
3				
4				
5				
6				

$\Rightarrow$

A	B	C	D
?	?	+A1+B1	
		+A2+B2	+B2+C2
		+A3+B3	+B3+C3
		+A4+B4	+B4+C4
		+A5+B5	+B5+C5
		+A6+B6	+B6+C6

A képernyőn nem a definíciók lesznek láthatók, hanem azok konkrét értékei, A1 és B1 mező tartalmától függően. Ha azonban az F2(Edit) billentyűvel változtatni szeretnénk egy cella tartalmán, akkor a fenti kifejezések kerülnek szerkesztésre a legfelső sorba.

A matematikai kifejezésekben lévő hivatkozások változásának az oka az, hogy a kifejezések beírásakor a cellákra történő hivatkozások relatívak, azaz a C1-ben lévő +A1+B1 kifejezés azt jelenti, hogy a cellába helyezze el a vele azonos sorban, de tőle balra lévő második és a tőle balra lévő első oszlopban lévő számok összegét. A másolás ezt a relatív hivatkozást csúsztatja el.

Szükség lehet arra is, hogy a másolás során az egyik operandus ne változzon, pl. ha bizonyos számokat egy konstanssal kell elosztani, amelyet egy korábbi számolás eredményeként kaptunk. Erre szolgál az ún. abszolút hivatkozás. Ha pl. a C2 mezőbe a **+B2/B1** kifejezést írjuk és ezt másoljuk a C oszlopban lefelé, akkor a másolatok mindig az előző oszlop (B) egymás feletti számainak hányadosát adják. Ha azonban minden B oszlopban lévő számot B1-gyel akarunk osztani, akkor a **+B2/$B$1** kifejezést kell írni a C2 mezőbe. Másolás után az eredmény a várt lesz, pl. a C4 mező tartalma **+B4/$B$1**-re változik.

A **$B$1** kifejezéssel a sor- és oszlophivatkozást is abszolúttá tettük, azaz a kifejezés értéke akkor sem változik, ha oszlop helyett sor szerint másolunk. Az oszlop szerinti másolásnál elégséges a hivatkozást a sor szerint abszolúttá tenni, azaz a **+B2/B$1** kifejezést lefelé másolva a C oszlopban ugyanazt kapjuk mint az előbb.

A mezőkijelöléshez használható "mutogatós" eljárásnál az abszolút hivatkozást jelölő $ karaktert az F4(Abs) billentyű lenyomásával tudjuk elhelyezni egy kifejezésben. Ezt akkor tehetjük meg, amikor a hivatkozandó mezőn állunk. Az F4-et többször lenyomva beállíthatjuk a kívánt formájú abszolút hivatkozást (sor+oszlop, csak sor, csak oszlop), illetve fel is szabadíthatjuk azt.

A másik fontos parancs a MoveCell, melynek formátuma azonos a CopyCell parancséval. Eredményeként az eredeti terület üresen marad, viszont az elmozgatott terület megtartja a mezőkkel való korábbi összefüggéseit. Ez akkor lehet fontos, ha valamely korábbi eredményt a jobb átláthatóság vagy más ok miatt át akarunk helyezni.

A tartományokkal való műveletek néha megkövetelik, hogy a kijelölt részek sarkai közt mozogjunk. Ezt ugyanazzal a pont (**.**) billentyű lenyomásával tehetjük meg, mint amivel lehorgonyoztuk a kijelölendő terület egyik sarkát. A pont többszöri lenyomása körbe-körbe visz a négy sarok között.

Másoláskor kétféle meglepő jelenség következhet be. Az egyik az, ha a várt eredmény helyett az **ERR** felirat jelenik meg a mezőben. Ekkor olyan művelet elvégzésére adtunk utasítást, amely értelmetlen, pl. nullával osztottunk. Vizsgáljuk át a képletünket, hogy utalásként tartalmaz-e olyan mezőt, amely a hibát okozhatja. A másik lehetséges eset az, hogy az eredmény helyett a teljes mezőszélességben csak csillagokat látunk. Ettől a kifejezés még tökéletes, csak az eredmény az adott formátumban nem jelezhető ki. Erre két megoldás van. Az egyik a /WSheet ColWidth Set parancsok kiadása után az oszlopnak a kurzormozgató billentyűkkel történő szélesítése, vagy a /Range Format parancsok kiadása után a megfelelő kijelzési formátum megválasztása. A formátumváltásnál a kijelezendő tizedesek számának megadása is szükséges. Legtöbbször a két eljárás együtt ad megfelelő eredményt.

A legfontosabb menüpontok

E rövid ismertető utolsó része azokat a menüből elérhető parancsokat ismerteti, amelyekre a fizikai-kémiai laboratóriumi gyakorlatokon történő kiértékelés során a legnagyobb szükség van.

- Adatokat menteni a /File Store utasítással és az adatállomány nevének megadásával lehet. Kiterjesztést nem szükséges megadni, ekkor a program automatikusan a **.WKS** kiterjesztést alkalmazza. Jó szokás néhány percenként elmenteni a munkalapot, mert ez nem igényel sok időt és nem vész kárba a felhasználó munkája egy esetleges technikai hiba miatt.
- Elmentett adatok visszaolvasása a /File Retrieve utasítás után a fájlnév kiválasztásával érhető el. Szükség esetén lehetséges váltani az aktuális könyvtárat.
- Egyszerű ASCII adatállományban lévő adatok behívása a /File Import Values utasítással hajtható végre.
- Nyomtatás a /PrintTo parancs után fájlba és nyomtatóra is lehetséges. Így tudjuk "exportálni" adatainkat ASCII adatállományokba, amiket bármilyen más program is "megért".

- Egyenes illesztése adatokra a /Data Regress parancs kiadásával végezhető. Az illesztés végrehajtása előtt értelemszerűen ki kell jelölni a független és függő adatok tartományát (X- és Y-változók), valamint a munkalap azon területét, ahová az illesztés eredményeit a program kiírja.
- A /Data Sort parancs segítségével sorba rendezhetők az adatok a kijelölt mezőtartományon belül (D-Range) növekvő (A:ascending) vagy csökkenő (D:descending) irányba, egy elsődleges rendezési elv (P(1): primary key) és/vagy egy másodlagos rendezési elv (S(2): secondary key) szerint. Például, ha személyek adatait akarjuk sorrendbe állítani, akkor elsődleges rendezési elv — más néven kulcs — lehet a név, másodlagos kulcs pedig a születési idő. Ez utóbbi természetesen csak akkor számít, ha a nevek megegyeznek.

  A rendezési elv(ek) megadásakor az adott oszlopot jelöljük meg, azaz az együtt mozgatandó elemeknek egy sorban kell lenniük. A számokat természetesen nagyságuk, a szövegkonstansokat pedig az ABC (pontosabban az ún. ASCII-kód) szerint rendezi a program.
- Megrajzolandó görbéket a /Graphics parancs kiadása után tudunk készíteni. Egy grafikonon hat görbe rajzolható meg tetszőleges, egymástól független módon, *de mindegyik Y-tartományt pontosan ugyanahhoz az X-tartományhoz rendeli a program!*
- Végül a legfontosabb parancs az F1(Help) funkcióbillentyű. Az itt leírtak megértése és kis gyakorlás után a program lehetőségei szinte teljes mélységükben megérthetők és kihasználhatók, minimális angol tudással és maximális tanulnivágyással.

II.F2. függelék
Számítógépes ábrakészítés

Ugyanazok a követelmények a számítógépes ábrákkal szemben, mint a kézzel, milliméterpapíron elkészített ábrákkal kapcsolatban:

– Minden mért adat (vagy azok transzformáltjai) szerepeljen az ábrán.
– Legyen megfelelő — szükség esetén mértékegységgel ellátott — címe mind a tengelyeknek, mind az ábrának. A feliratok mind szakmai, mind nyelvtani szempontból legyenek helyesek. Név és dátum szerepeljen az ábrán.
– A tengelyek beosztásának és a címkefeliratoknak olyanoknak kell lenniük, hogy az adatok könnyen ábrázolhatók és visszaolvashatók legyenek. Ezt az elvet mindig a konkrét feladatra kell alkalmazni, pl. egyenesillesztés esetén néha szükséges, hogy a tengelymetszet akkor is rajta legyen az ábrán, ha kívül esik a mért adatok tartományán.
– Görbeillesztés esetén az ábrának tartalmaznia kell mind az illesztett, mind az illesztésből kihagyott adatokat (megkülönböztetett jelzéssel!), valamint az illesztett görbét is, az illesztett paraméterek értékével együtt.
– Több görbe együttes ábrázolása esetén az egyes görbék vagy adatsorok legyenek világosan elkülöníthetők.

Természetesen lehetnek egyéb követelmények is a konkrét feladattól függően. Ritkán a követelmények egyike-másika nem teljesíthető 100%-osan, de az esetek túlnyomó többségében a fentiek betartása elegendő hibátlan ábrák készítéséhez.

Az eddigiekből is kitűnik, hogy az ábrakészítő programok felületes ismerete nem elég. *Nem lehet elfogadni kész ábraként, amit egy program az adatok bevitele után az alapbeállításaival felrajzol a képernyőre, hanem olyan mélységig kell megtanulni az alkalmazott program kezelését, hogy a fenti követelmények teljesíthetők legyenek!* A tudományos életben ezt fokozottan kell hangsúlyozni, mert a legtöbb kereskedelemben lévő program (főleg a táblázatkezelők) a közgazdasági és prezentációs célokra készített ábrákhoz igazítja az alapbeállításokat, és nem a tudományos életben fokozottabban megkövetelt pontosság és teljesség igényéhez.

Az alábbiakban egy példán keresztül bemutatjuk azokat a jellemző hibákat, amiket a leggyakrabban szoktak elkövetni számítógépes ábrakészítés során. Az F2.1 és F2.2 ábrák ugyanarra az adatsorra történő egyenesillesztést illusztrálják. Az F2.1 ábra teljes mértékben megfelel a fentebb részletezett követelményeknek, míg az F2.2 ábra a — tapasztalatunk szerint — leggyakrabban előforduló hibákat mutatja. Ezek könnyen elkerülhetők a használt program megfelelő mélységű ismeretével. A jelen függelék a két ábra összevetésével segíteni igyekszik a következő tipikus hibák és hiányosságok elkerülését:

Automatikus pontösszekötés. Majdnem minden program alapbeállítása az, hogy a bevitt pontokat valamilyen szimbólummal jelöli és azokat egyenes szakaszokkal köti össze. Az összekötésnek a legtöbb esetben nincs értelme, csak a "szem vezetésére" szokták használni tendenciák bemutatására közgazdasági grafikonokon. Tudományos ábrákon vonallal illesztett görbét szokás jelölni, így a mért pontok összekötése félrevezető. Ráadásul, értelmezhetetlen ábrákhoz vezethet, ha az adatok nincsenek szigorúan növekvő vagy csökkenő sorrendben, pl. az F2.2 ábrán egyetlen nem sorban lévő pont egy felesleges vonalat ad.

Egyenetlen tengelybeosztáshoz vezet sok program azon tulajdonsága, hogy a tengelyek minimális és maximális értékét az adatsorokból kapható minimális és maximális értékhez rendeli. Az F2.2 ábrán az Y-tengely beosztása rossz, mert a 13–120 tartományt nem lehet jól felosztani tíz részre. Ráadásul a beosztás-feliratok pontatlanok (csak egész értékre vannak megadva). Emiatt nemcsak kényelmetlen a visszaolvasás, hanem hibás is, azonos hosszúságú tartományokhoz eltérő értékek tartoznak (120–109≠109–99)! A felhasználónak tudnia kell, hogyan lehet beállítani a tengelyek minimális és maximális értékeit, a beosztás sűrűségét és a beosztások helyét jelző számok kiiratási formátumát!

Automatikus tengelyválasztás miatt rossz beosztás, értelmetlen vagy hiányzó címek és beosztásfeliratok lehetnek a tengelyeken. Az F2.2 ábrán hiányoznak az X-tengely beosztásainak feliratai, a semmitmondó automatikus tengelycímek az adatállomány nevéből és a felhasznált oszlopok sorszámából adódnak, a legszélső adatok pedig szinte "lelógnak" az ábráról.

Semmitmondó főcím nehezíti az ábra megértését, főleg ha az értelmezés és a készítés között sok idő telik el. Sok program alapértelmezése, hogy főcímként a grafikus beállításokat tartalmazó állomány nevét definiálja.

Név, cím vagy dátum hiánya szintén bosszantó információvesztés lehet. Példánkban a rossz ábrán a dátum hiányzik.

Rossz pozícionálása az ábra valamelyik részének jobbik esetben csak komikus, rosszabb esetben információvesztéshez vezethet. Esetünkben a magyarázó blokk annyira rossz helyen van az F2.2 ábrán, hogy a fele lemaradt.

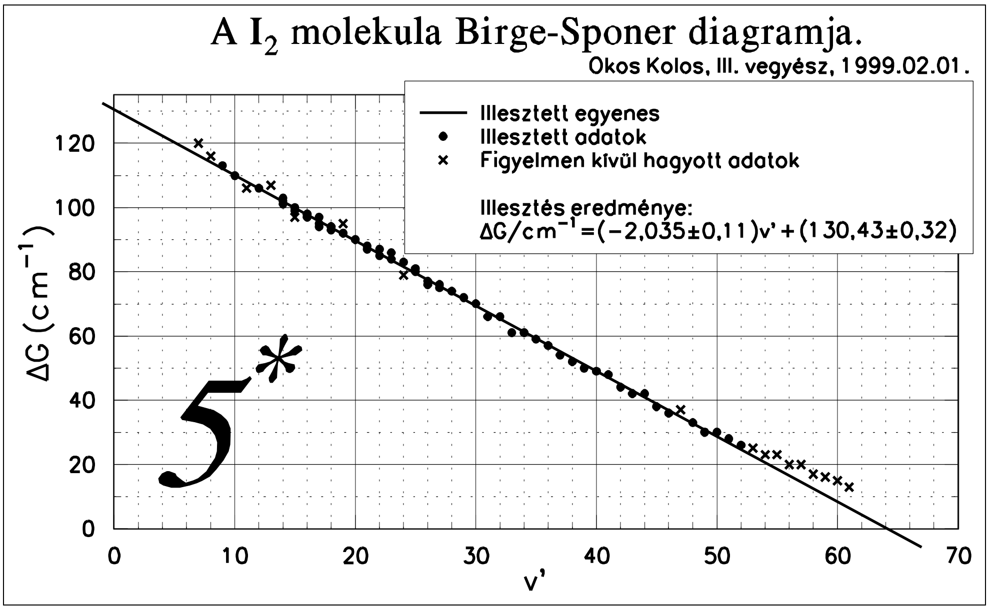
Az **automatikus magyarázó blokk** (angolul legend) általában semmit sem mond. Vagy egyáltalán ne használjunk ilyen blokkot, vagy pontosan töltsük ki! A magyarázó blokknak akkor van értelme, ha egy ábrán több görbét tüntetünk fel és rövid utalásokkal akarjuk segíteni az ábra megértését.

A **rácsozás**t nagyon gondosan kell beállítani! Nem segíti az ábra olvasását a túl sűrű rácsozat, mert az szinte elfedi a görbét. A ritka rácsozat sem jó, mivel a függvényértékeket nagyon nehéz visszaolvasni ilyen esetekben.

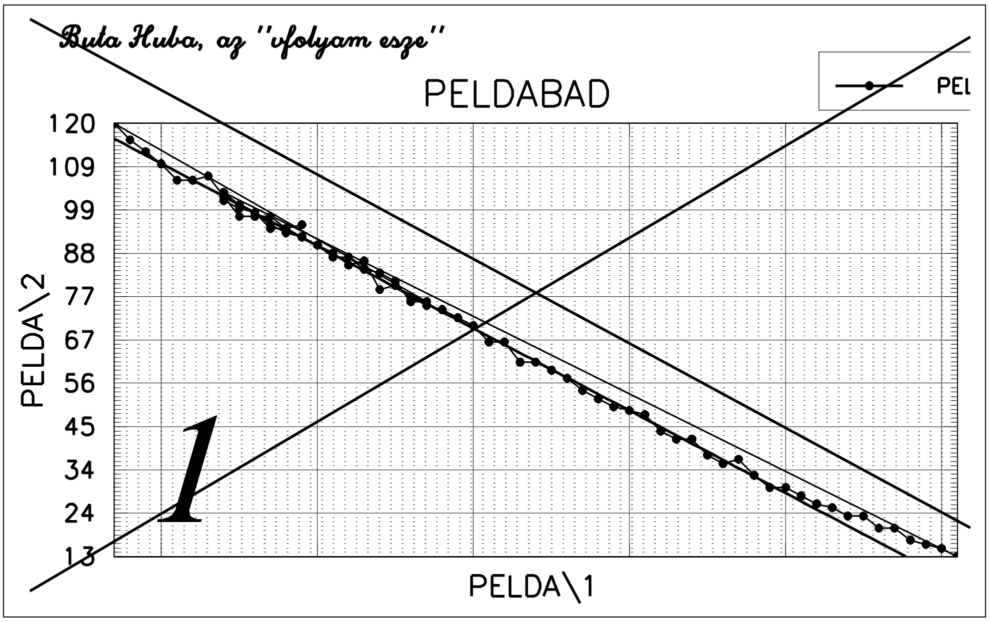
Nem megfelelő betűtípus és betűnagyság használata jobbik esetben csak csúnya és komikus feliratokhoz vezet, rosszabik esetben félreértésre is alkalmat adhat. Az F2.2 ábrán a név értelmetlenül csicsás és nem ékezetes betűkkel készült.

Ábrákon általában érdemes egyszerű vonalvezetésű és vastagabb betűtípusokat használni (pl. Arial, Helvetica, Swiss, stb. típusok).

Illesztésből kihagyott pontok vagy nem szerepelnek, vagy jelöletlenek. Ha az illesztésnél figyelmen kívül hagyott pontokat nem tüntetjük fel az ábrán, akkor információt vesztünk a mérések valódi pontosságáról. Ha a rossz pontokat az illesztettekkel azonos módon jelöljük, akkor a számított adatok reprodukálása lesz nehéz.



F2.1 ábra: Egy kifogástalanul elkészített számítógépes ábra.



F2.2 ábra: A tipikus hibákat bemutató számítógépes ábra.

II.F3. függelék
Egyenes illesztése hibával terhelt adatokra

A fizikai-kémia gyakorlatokon a leggyakrabban használt kiértékelő módszer egy egyenes illesztése a mérési adatokra vagy azok transzformáltjaira. Ennek végrehajtásához a tanszék biztosít egy DOS alatt működő programot **EGYENES.EXE** néven, amely minden hallgatói számítógépen futtatható. A program a tanszék honlapról is letölthető. Használatával — a matematikai háttér pontos ismerete nélkül — minden olyan statisztikai adat kiszámítható, amelyet a gyakorlatokon megkövetelnek.

Amint azt a bevezetésben is említettük, az emelt szintű gyakorlatok egyik célja, hogy a hallgatókat "rászoktassa" táblázatkezelők és egyéb, tudományos célú értékelő programok használatára. Az ilyen programok tudása sokban különbözik, nem mindegyikük számítja automatikusan a szükséges paramétereket (pl. **AS-EASY-AS...** 5.50c verziója nem adja meg a tengelymetszet szórását). A hiányzó, de megkövetelt adatokat a hallgatóknak maguknak kell kiszámítaniuk, amelyre legalább két módszer lehetséges:

1. Az illesztendő adatokat ASCII-fájlba exportáljuk és az egyenes illesztését a már említett **EGYENES.EXE** program segítségével hajtjuk végre.
2. Az előzőnél sokszor kényelmesebb és/vagy hatékonyabb megoldás, ha a szükséges adatokat a használt programon belül, a megfelelő matematikai képletek alapján számítjuk ki.

A jelen függelék két alfejezete az utóbbi megoldáshoz nyújt segítséget. Az első fejezet az egyenesillesztés matematikai alapjait, valamint a statisztikai adatok számításához szükséges képleteket tartalmazza. Általános tapasztalat, hogy a statisztikai adatok értelmezése sok gondot okoz, ezért a lényegesebb paraméterek értelmezését megadjuk a képletek után.

A második alfejezet egy példán keresztül bemutatja, hogyan lehet alkalmazni az első alfejezetben és a II.F2. függelékben leírtakat az **AS-EASY-AS...**-ben ahhoz, hogy igényes ábrát tudjunk készíteni korrekt számításokkal.

II.F3.1 Egyenes illesztésének matematikai alapjai

Adott n darab (x_i , y_i) adatpár, amelyek általában mérési adatok vagy azok transzformáltjai. Keressük annak az y=ax vagy y=ax+b formában megadható egyenesnek a paramétereit (a és b), amely a legjobban illeszkedik az adatokra, vagyis az illesztett és a modell alapján számított adatok eltéréseinek négyzetösszege ($S_q \equiv S^2$) minimális. A négyzetösszegfüggvényt az a és b paraméterek függvényeként definiáljuk úgy, hogy az x_i értékeket teljesen pontosnak tekintjük és az összes hibát az y_i értékeihez rendeljük:

$$S_q(a) = \sum_{i=1}^{n} (y_i - ax_i)^2, \text{ ha a modell } y = ax \quad \text{és}$$
$$S_q(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2, \text{ ha a modell } y = ax + b. \tag{F3.1}$$

Az a, b paraméterek értékeinek kiszámításához azt használjuk fel, hogy a négyzetösszegfüggvény differenciálhányadosának (vagy parciális differenciálhányadosainak) értéke zéró a minimum helyén. Ezt fejezi ki az origón átmenő egyenes esetén a

$$\frac{\mathrm{d}S_q(a)}{\mathrm{d}a} = 0 = \sum_{i=1}^{n} -2(y_i - ax_i)x_i = 2a\sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i y_i\,, \tag{F3.2}$$

összefüggés, illetve az $y=ax+b$ modell esetén egy bonyolultabb, a paraméterekre nézve lineáris egyenletrendszer, ami a

$$\begin{aligned} \frac{\partial S_q(a,b)}{\partial a} &= 0 = \sum_{i=1}^{n} -2(y_i - ax_i - b)x_i = 2a\sum_{i=1}^{n} x_i^2 + 2b\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} x_i y_i \\ \frac{\partial S_q(a,b)}{\partial b} &= 0 = \sum_{i=1}^{n} -2(y_i - ax_i - b) = 2a\sum_{i=1}^{n} x_i + 2nb - 2\sum_{i=1}^{n} y_i \end{aligned} \tag{F3.3}$$

formában adható meg. Az (F3.2) egyenlet vagy az (F3.3) lineáris egyenletrendszer megoldása szolgáltatja a meredekség és az esetleges tengelymetszet értékét. A matematikai statisztika segítségével az illesztett adatok, illetve az illeszkedés mértékének jellemzésére szolgáló statisztikai adatok is számíthatók. A részletes levezetéseket elhagyva, az F3.1 táblázat adja meg a számításokhoz szükséges matematikai képleteket. A fogalmak nevét angolul is megadjuk, mivel a legtöbb felhasználói program ezen a nyelven kommunikál (ha többféle elnevezés is használatos, akkor a leggyakoribbat adtuk meg).

F3.1 táblázat: Az egyenesillesztés matematikai képletei.

fogalom	jelölés	definíció matematikai képlete	
magyarul *angolul*		*y*=*ax* modell	*y*=*ax*+*b* modell
meredekség *slope*	a	$\dfrac{\sum_{i=1}^{n} x_i y_i}{\left(\sum_{i=1}^{n} x_i\right)^2}$	$\dfrac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$

fogalom magyarul *angolul*	je- lölés	definíció matematikai képlete $y{=}ax$ modell	 $y{=}ax{+}b$ modell
tengelymetszet *intercept*	b	—	$\dfrac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right)-\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}$
négyzetösszeg *sum of squared residuals*	S	$\sqrt{\sum_{i=1}^{n}(y_i-ax_i)^2}$	$\sqrt{\sum_{i=1}^{n}(y_i-ax_i-b)^2}$
átlagos eltérés *standard deviation of data*	σ	$\sqrt{\dfrac{S^2}{n-1}}$	$\sqrt{\dfrac{S^2}{n-2}}$
meredekség standard hibája *standard error of the slope*	S_a	$\sqrt{\dfrac{n}{n-1}\dfrac{S^2}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}}$	$\sqrt{\dfrac{n}{n-2}\dfrac{S^2}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}}$
tengelymetszet standard hibája *standard error of the intercept*	S_b	—	$\sqrt{\dfrac{\sum_{i=1}^{n} x_i^2}{n-2}\dfrac{S^2}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}}$
meredekség standard szórása *standard deviation of the slope*	σ_a	$\sqrt{n}\,S_a$	
tengelymetszet standard szórása *standard deviation of the intercept*	σ_b	—	$\sqrt{n}\,S_b$
korrelációs együttható *correlation coefficient*	r	$\dfrac{n\sum_{i=1}^{n} x_i y_i-\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\left(n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2-\left(\sum_{i=1}^{n} y_i\right)^2\right)}$	

fogalom magyarul *angolul*	jelölés	definíció matematikai képlete $y=ax$ modell	 $y=ax+b$ modell
a teljes variancia illesztéssel magyarázott része *coefficient of determination*	CoD	$(1-n)\dfrac{S^2}{n\sum\limits_{i=1}^{n} y_i^2-\left(\sum\limits_{i=1}^{n} y_i\right)^2}$	
az összes észlelés illesztéssel magyarázott része *R-squared*	R	$1-\dfrac{S^2}{\sum\limits_{i=1}^{n} y_i^2}$	
kritikus t-érték x%-os megbízhatósági szinten *critical t-value at x% confidence level*	$t_{x\%}$	Student-eloszlás(n–1, 0,01$\cdot x$)	Student-eloszlás(n–2, 0,01$\cdot x$)
variancia-kovariancia mátrix *variance-covariance matrix*	$\boldsymbol{V}$	—	$\begin{pmatrix} \dfrac{n}{n\sum\limits_{i=1}^{n} x_i^2-\left(\sum\limits_{i=1}^{n} x_i\right)^2} & \dfrac{-\sum\limits_{i=1}^{n} x_i}{n\sum\limits_{i=1}^{n} x_i^2-\left(\sum\limits_{i=1}^{n} x_i\right)^2} \\ \dfrac{-\sum\limits_{i=1}^{n} x_i}{n\sum\limits_{i=1}^{n} x_i^2-\left(\sum\limits_{i=1}^{n} x_i\right)^2} & \dfrac{\sum\limits_{i=1}^{n} x_i^2}{n\sum\limits_{i=1}^{n} x_i^2-\left(\sum\limits_{i=1}^{n} x_i\right)^2} \end{pmatrix}$
korrelációs mátrix *correlation matrix*	$\boldsymbol{C}$	—	$\begin{pmatrix} 1 & \dfrac{-\sum\limits_{i=1}^{n} x_i}{\sqrt{n\sum\limits_{i=1}^{n} x_i^2}} \\ \dfrac{-\sum\limits_{i=1}^{n} x_i}{\sqrt{n\sum\limits_{i=1}^{n} x_i^2}} & 1 \end{pmatrix}$

Megjegyzések:

1. Az F3.1 táblázat képletei első pillantásra bonyolultnak tűnhetnek. A számítások jóval egyszerűbbé válnak, ha az összefüggésekben bevezetjük a következő jelöléseket: $\mathrm{d}_1 = n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2$, $\mathrm{d}_2 = n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2$ és $\mathrm{d}_3 = n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)$.
2. A variancia-kovariancia- és a korrelációs mátrix között elemenként egyértelmű összefüggés van, amit a $C_{ij} = V_{ij} \Big/ \sqrt{V_{ii} V_{jj}}$ képlet ad meg.

A statisztikai paraméterek használatában a leggyakoribb félreértelmezések a következők:

1. *Egy illesztett paraméter* <u>*standard hibája*</u> *és a* <u>*standard szórása*</u> *nem ugyanaz!* Az illesztett paraméter értékének bizonytalanságát *csakis a szórás* jellemezheti! Ennek belátásához gondoljuk végig a következőket:
   a) Az nyilvánvaló, hogy egy illesztett paraméter értéke nem lehet pontosabb, ha minden illesztendő (x_i, y_i) adatpárt többször adunk meg az egyenesillesztő programnak. Csak olyan statisztikai paraméter jellemezheti egy paraméter bizonytalanságát, melynek számértéke nem (vagy csak elhanyagolhatóan) függ attól, hogy hányszor ismétlünk meg adatpárokat az illesztésben.
   b) Elemezzük a standard hibák matematikai kifejezéseit (az F3.1 táblázat S_a-t és S_b-t definiáló sorai). Ha van egy konkrét adatsorunk, akkor a képletek alapján könnyen kiszámolható pl. a meredekség hibája. Jelöljük ezt az értéket S_a^1-val.
   c) Ismételjük meg az illesztést úgy, hogy minden adatpárt k-szor adunk meg, vagyis az adatok száma $k \cdot n$ lesz. Így is számítsuk ki a meredekség standard hibáját, amit jelöljünk S_a^k-val.
   d) A definiáló képletek alapján S_a^k a következő módon fejezhető ki S_a^1 segítségével:

$$S_a^k = S_a^1 \sqrt{\frac{n-2}{kn-2}} \quad \rightarrow \quad \lim_{k\to\infty} S_a^k = \lim_{k\to\infty} \frac{S_a^1}{\sqrt{k\dfrac{n}{n-2} - \dfrac{2}{n-2}}} = 0 \; ,$$

   vagyis a standard hiba számított értéke tetszőlegesen kicsire csökkenthető akár úgy is, hogy ugyanazokat a mérési adatokat nagyon sokszor adjuk meg az illesztéshez!
   e) A fenti számításokat nemcsak a hibákra, hanem a standard szórásokra is végre lehet hajtani. Analóg jelölésekkel élve a következő összefüggésekhez jutunk:

$$\sigma_a^k = \sigma_a^1 \sqrt{\frac{k(n-2)}{kn-2}} \quad \rightarrow \quad \lim_{k\to\infty} \sigma_a^k = \lim_{k\to\infty} \sigma_a^1 \sqrt{\frac{n-2}{n-\frac{2}{k}}} = \sigma_a^1 \sqrt{\frac{n-2}{n}} \; ,$$

vagyis a standard szórás értéke a k növelésével nem változik számottevően, csak az eredeti érték $\sqrt{(n-2)/n}$-szeresére csökken, ami n=3 esetén 42%-os változást jelent (ez a legrosszabb eset!), de pl. n=10 esetén már csak 10%-os csökkenést.

Az eddigiekből következik, hogy a szórás jellemzi az illesztett paraméter megbízhatóságát, a standard hiba pedig a paraméter szórásának a szórása, ami azt mondja meg, hogy a standard szórás számértéke (és *nem* a paraméteré!) milyen mértékben megbízható!

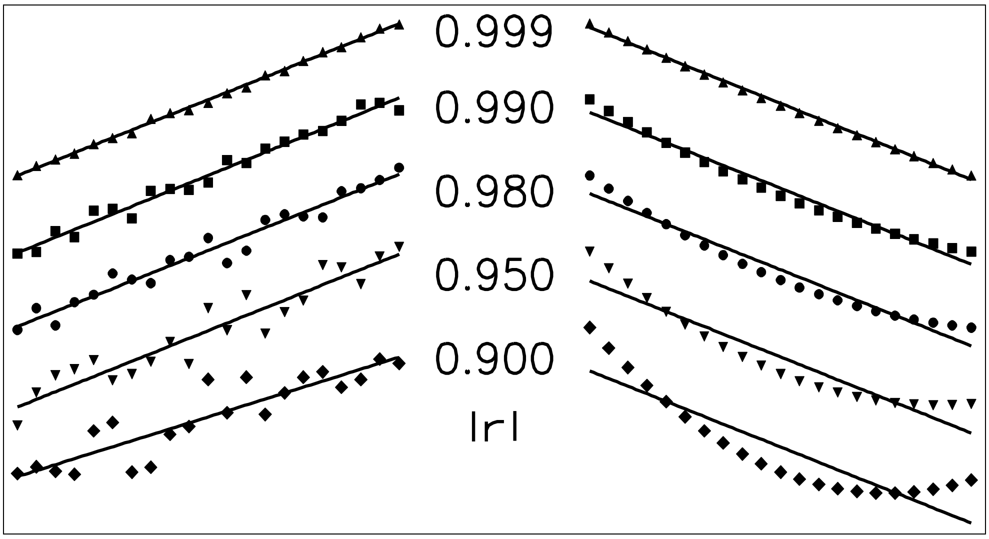
Sok kereskedelmi program keveri a standard szórás és standard hiba fogalmát és szórásként adják meg a hiba számértékét. A fenti levezetések alapján egyszerűen el lehet dönteni, hogy a használt program mit számol. Először végezzük el az egyenesillesztést egy adatsorral (legalább 20 adat!). Ezután minden illesztendő adatpárt adjunk meg négyszer és végezzük el újra az illesztést. Ekkor — az F3.1 táblázat képleteiből és a fenti meggondolásokból következően — a standard hibának kb. a felére kell csökkennie, míg a standard szórás nem változhat lényegesen. Ha a program hibát ad meg a szórásértékként, akkor ezt meg kell szorozni az adatok számának négyzetgyökével ahhoz, hogy megkapjuk a valódi szórást.

2. Az illeszkedés mértékének jellemzésére az egyik leggyakrabban használt statisztikai paraméter a korreláció. *Ez a szó önmagában nem egyértelmű* (sajnos szinte csak így használják!), ezért sokan keverik a *korrelációs együttható* (r) és a *korrelációs mátrix* nem diagonális elemének (C_{ij}) jelentését. Mindkettő egy normált érték –1 és 1 között, de összesen ennyi a közös bennük. Megnézve az F3.1 táblázatot kitűnik, hogy ez a két adat teljesen más képlettel számolható, így más fogalmat is takar:

   a) Az r értéke függ az x_i és y_i adatoktól is, míg a C_{ij} érték *csak* az x_i adatoktól.

   b) Az r azt fejezi ki, hogy az egyenes — mint modell — milyen mértékben képes leírni az (x_i,y_i) adatpárok összefüggését, így az illeszkedés mértékének egyik mérőszáma. Másképp fogalmazva r azt mondja meg, hogy milyen erős a lineáris kapcsolat a függő és független változók adatai között. Minél közelebb van r abszolút értéke egyhez, annál valószínűbb a lineáris kapcsolat. A *közel* szót komolyan kell venni! Szemléletesen pl. a 0,95-ös érték is közel van 1-hez, de ha r értéke ennyi, akkor a lineáris kapcsolat már nehezen valószínűsíthető, vagy az adatok nagy szórása, vagy azok menete miatt (ld. az F3.1 ábrát, valamint az alfejezet többi példáját). Általánosságban az mondható, hogy 10–20 mért adat és három értékes jegyre pontos mérések esetén

0,995 vagy nagyobb érték valószínűsíti a lineáris összefüggést, míg 0,98 alatti érték már pontatlanabb mérések esetén is kérdőjelessé teszi azt. Jóval több adat (több, mint 100) esetén a fenti határok 1–2%-kal csökkenhetnek.



F3.1 ábra: Az illesztett adatok normális eloszlású hibája (bal oldal) vagy görbülete (jobb oldal), valamint a különböző korrelációs együttható (r) értékek közötti összefüggés szemléltetése.

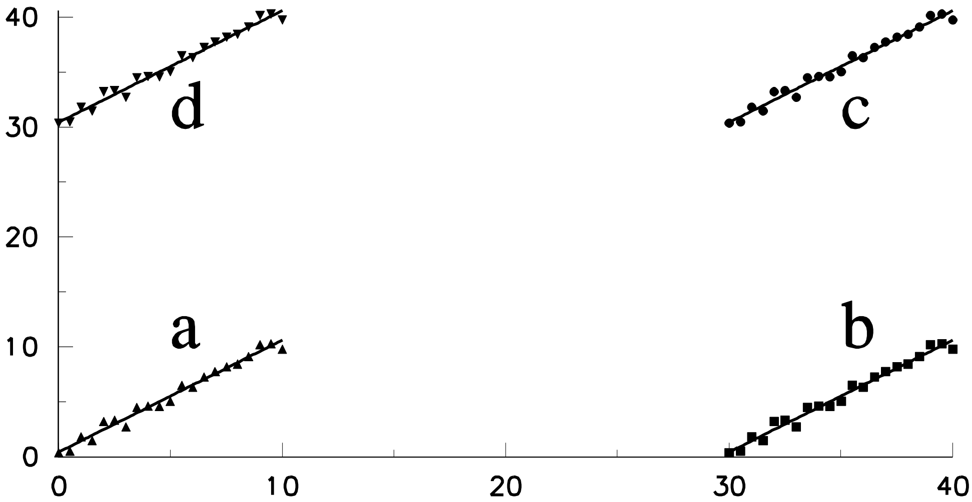
Az F3.1 ábra jól szemlélteti a leírtakat. Az ábrázolt tíz adatsor x_i értékei és a függő változók tartománya ugyanaz, csak az y_i értékekben változtattuk a véletlen hibák nagyságát vagy a görbék menetének erősségét a kívánt r-nek megfelelően.

c) A b. ponttól eltérően, C_{ij} azt fejezi ki, hogy a meredekség és a tengelymetszet mennyire függetlenek egymástól. Az könnyen megérthető, hogy ez a paraméter csak az x_i értékektől függ. Minél távolabb van a tengelymetszet helye a mért x_i pontok által behatárolt tartománytól, annál bizonytalanabb a tengelymetszet, hiszen a meredekség kis változása a tengelymetszet nagy változását vonja maga után. Minél távolabb van a mérési tartomány a tengelymetszettől, annál inkább *nem* a mért adatok, hanem a meredekség bizonytalansága szabja meg b értékét, ezért a két paraméter egyre inkább korrelál. A teljes korrelációnál C_{ij} abszolút értéke 1, így — a korrelációs együtthatóval ellentétben! — 1-től minél inkább eltérő érték biztosítja azt, hogy mind az a, mind a b értéke biztonsággal meghatározható legyen.

3. Mind az R, mind a CoD 0 és 1 közé normált érték. A konkrét értékek értelmezéséről ugyanaz mondható el, mint a korrelációs együttható esetében. R azt fejezi

ki, hogy az y_i adatok *értékét* mennyire képes (relatív értelemben) visszaadni az illesztett egyenes. Ettől eltérően, CoD azt mondja meg, hogy az y_i adatok *változásának mértékét* mennyire képes az illesztett egyenes leírni. Ha y_i-k számértéke a változásokhoz képest nagy, akkor R értéke mindig közel lesz egyhez — a legrosszabb illeszkedés esetén is! —, míg CoD számértéke továbbra is mérvadó információ. Ennek ellenére az R értéket sokkal többet használják az irodalomban.

<table>
<tr><th></th><th>$a \pm S_a$</th><th>$b \pm S_b$</th><th>σ</th><th>r</th><th>S^2</th><th>CoD</th><th>R</th><th>C_{ij}</th></tr>
<tr><td>a</td><td rowspan="4">1,018 ±0,033</td><td>0,43±0,19</td><td rowspan="4">0,461949</td><td rowspan="4">0,989994</td><td rowspan="4">4,05455</td><td rowspan="4">0,980087</td><td rowspan="2">0,99519</td><td>−0,855399</td></tr>
<tr><td>b</td><td>−30,1±1,2</td><td rowspan="2">−0,996279</td></tr>
<tr><td>c</td><td>−0,1±1,2</td><td rowspan="2">0,999848</td></tr>
<tr><td>d</td><td>30,43±0,19</td><td>−0,855399</td></tr>
</table>



F3.2 ábra: Normális eloszlású hibával terhelt négy adatsor az illesztett egyenesekkel, valamint a számított statisztikai paraméterek.

A tárgyalt paraméterek jelentését segít megérteni az F3.2 ábra. Az **a** adatsor ugyanaz, mint az F3.1 ábra bal oldalának r=0,99-es korrelációs együtthatójú adatsora. Mind a függő, mind a független változók értékeiben ~10 a különbség a legnagyobb és legkisebb adat között. A többi görbét úgy állítottuk elő, hogy az **a** adatsort eltoltuk az adattartomány háromszorosával, **b**-nél a vízszintes, **d**-nél a függőleges, **c**-nél pedig mindkét tengely irányában. Ezek a transzformációk az illeszkedés jóságát nem változtathatják meg. A számított statisztikai paraméterekből a következők láthatók:

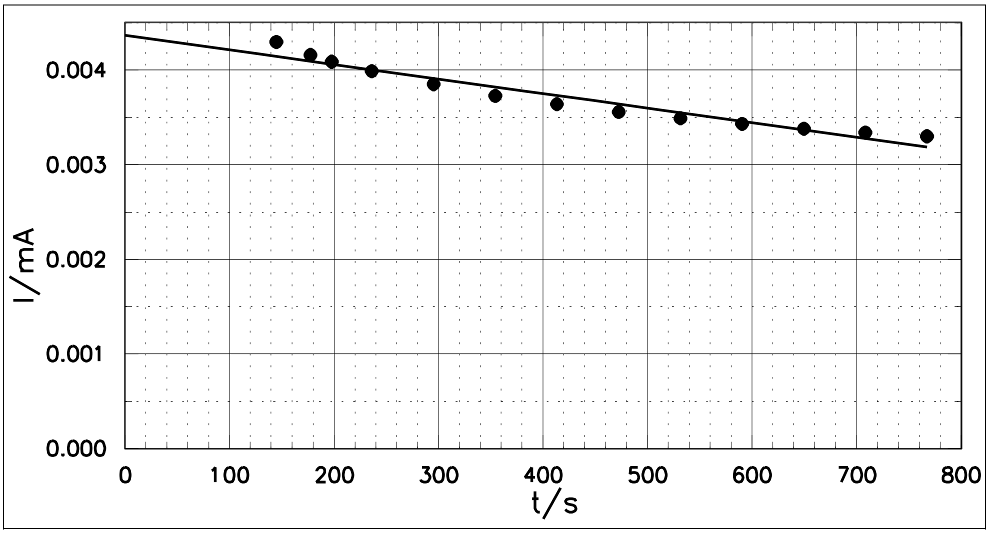
- Nem igényel különösebb magyarázatot az a tény, hogy a meredekséget (a), valamint a mért és számított adatok átlagos eltérését (σ) és ezek négyzetösszegét (S^2) a transzformációk nem változtatják meg, míg a tengelymetszet (b) és hibája más és más mind a négy esetben.
- Nem változik a korrelációs együttható (r) és a teljes variancia illesztéssel megmagyarázott része (CoD) sem, így ezek a paraméterek alkalmasak az illesztés jóságának jellemzésére. A két paraméter között gyakorlati szempontból csak az a különbség, hogy egy-egy kiugró pont esetén a CoD értéke sokkal jobban változik, mint r-é (könnyen kipróbálható konkrét adatokon).
- A teljes észlelés illesztéssel magyarázott részének (R) értékét a vízszintes tengely és az adatsor függőleges tartományának távolsága szabja meg alapvetően, így az illeszkedés mértékének jellemzésére kevésbé alkalmas.
- A korrelációs mátrix nem diagonális elemének (C_{ij}) abszolút értéke annál közelebb van egyhez, minél távolabb van az adatsor vízszintes tartománya a függőleges tengelytől. Szemléletesen ez azt jelenti, hogy minél távolabbra kell extrapolálni a tengelymetszet értékének meghatározásához, annál bizonytalanabb lesz annak értéke.

Egyenes illesztéséhez szinte kizárólagosan az F3.1 táblázatban megadott statisztikai paramétereket és képleteket használják. A kifejezések mindegyikét azzal a szigorú feltételezéssel vezették le, hogy az (x_i , y_i) adatok egymástól teljesen függetlenek. A képletekben az n–1 és n–2 részkifejezések valójában a szabadsági fokok számát jelentik (angolul *degree of freedom*, rövidítése DoF). Az illesztett adatok teljes függetlensége esetén DoF értéke valóban az illesztett adatok és az illesztett paraméterek számának különbségével adható meg. Ha azonban az illesztendő adatok között korreláció van (pl. az ismételt adatok között a korreláció teljes!), a levezetett képletek akkor is igazak, de a szabadsági fokok száma csökken (növekvő adatkorrelációval egyre drasztikusabban). A korreláció figyelembevétele matematikailag bonyolult, "nehezen emészthető", számítógépek nélkül gyakorlatilag alkalmazhatatlan eljárásokkal lehetséges, ezért szinte soha nem használják ezeket, még a kutatásban sem.

A valóságban a mérési adataink sohasem teljesen függetlenek. Például, minden mért mintát ugyanabból a törzsoldatból készítünk, ugyanazzal a készülékkel vizsgálunk, ugyanazzal a térfogatmérő eszközzel mérjük ki, stb. Emiatt a számított statisztikai paraméterek úgy kezelendők, mint ideális határértékek, amelyeknél a valóság csak rosszabb lehet! Ennek akkor van különös jelentősége, amikor egy paraméter fontosságáról vagy pontosságáról éppen szórásának számértéke alapján akarunk dönteni. Ha az illesztett adatpárok közötti korrelációkról nem tudunk mit mondani, akkor érdemes a standard szórás (vagy standard hiba) számított értékénél 2–3-szor nagyobb adatot elfogadni reális értékként.

Egy másik hasznos tanács, hogy soha ne döntsünk egyetlen statisztikai paraméter értéke alapján. Mindig számítsunk ki több paramétert és nézzük meg, hogy értékeikből külön-külön mire lehet következtetni. Ha a konklúziók egybehangzóak, akkor nincs baj, a következtetés nagy valószínűséggel helyes lesz. Ha a statisztikai paraméterek ellentmondanak egymásnak, akkor vagy a mérési adatokkal van gond (pl. nincs bennük információ az illesztett paraméterekre vonatkozóan), vagy az egyenes nem fogadható el, mint modell.

Az eddig leírtakat szemlélteti az F3.3 ábrán bemutatott egyenesillesztés. Az illesztendő pontok valós mérési adatok. Az összefüggés láthatóan nem írható le egy egyenessel. Az **EGYENES.EXE** a következő eredményeket számolta:



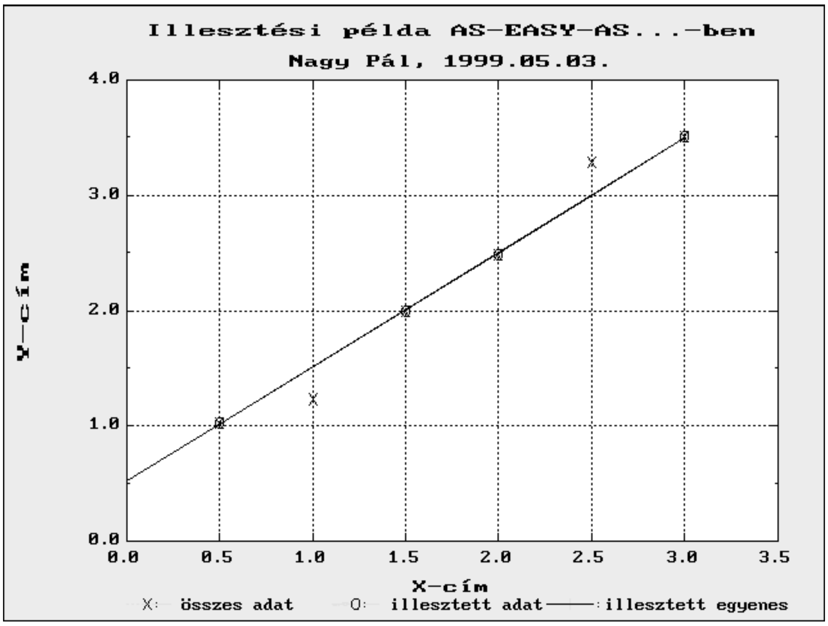
F3.3 ábra: Egy (elriasztó) példa az egyenesillesztés használatára.

```
                   Meredekség és annak standard hibája: -1,54439E-6±1,14798E-7
               Tengelymetszet és annak standard hibája:  4,37010E-3±5,41868E-5
          Korrigált empirikus szórás (átlagos eltérés):  8,42096E-5
                            Korrelációs együttható (r): -0,970930
                                 Eltérés négyzetösszeg:  7,80038E-8
A teljes variancia illesztéssel megmagyarázott része:  0,942705
Az összes észlelés illesztéssel megmagyarázott része:  0,999568
    Kritikus t-érték 95,0%-os megbízhatósági szinten:  2,201
    Variancia-kovariancia mátrix:                    Korrelációs mátrix:
a    1,85842E-6    -7,91545E-4                   a   1,000000  -0,902343
b   -7,91545E-4     4,14061E-1                   b  -0,902343   1,000000
```

A következők olvashatók ki ezekből az eredményekből:

1. A korrelációs együttható értéke világosan mutatja, hogy az összefüggés lineáris volta vagy az adatok megfelelő pontossága megkérdőjelezhető. Esetünkben a pontosság az F3.3 ábrán láthatóan megfelelő, ellenben a mért adatok között nem áll fenn lineáris kapcsolat.

2. A mérési tartomány elég közel van az origóhoz, így a tengelymetszet értéke kevéssé függ a meredekség hibájától. A gyakorlatban azt lehet mondani, hogy 0,95 alatti C_{ij} érték esetén mind az a, mind a b értékére elegendő független információ van a kísérleti adatokban.
3. Az y_i értékek változása körülbelül 20% az abszolút értékükhöz képest. Már ez is azt eredményezi, hogy R számított értéke gyakorlatilag egy lesz. A 20%-os változás bőven elég arra, hogy az F3.3 ábrára pillantva megállapítsuk: az adatsor nem írható le egy egyenessel a konkrét kísérleti pontosság mellett. Ezt a kézenfekvő tényt R számított értéke nem adja vissza, így alkalmatlan az illeszkedés jellemzésére (ennek ellenére, hogy az irodalomban gyakran alkalmazzák). Ezzel ellentétben CoD értéke jól mutatja, hogy a kísérleti adatokra nem illeszthető megalapozottan egyenes.



F3.4 ábra: Egyenesillesztést bemutató, **AS-EASY-AS...**-ben készült ábra.

II.F3.2 Egyenesillesztés az **AS-EASY-AS...** segítségével

A most bemutatandó példa azt illusztrálja, hogyan tehetünk eleget az ábrakészítés (II.F2. függelék) követelményeinek, valamint hogyan tudjuk kiszámítani azon statisztikai paramétereket, amelyeket az **AS-EASY-AS...** nem ad meg automatikusan.

Az F3.4 ábra az illesztés az egyik végeredménye. Látható, hogy az összesen hat adatpár közül négy szépen illeszkedik egy egyenesre, kettőt viszont célszerű a számításokból kihagyni. Ezért a valóban illesztett adatokat eltérően jelöltük. Az ábrán szerepel a számított egyenes is. A tengelyek ábrázolási tartományait úgy választottuk meg, hogy a tengelymetszet is látható legyen.

Érdemes megfigyelni, hogy szemmel viszonylag jó illeszkedésű egyenest kapnánk akkor is, ha az F3.4 ábra hat adata közül a két szélsőt hagynánk el. Az is érzékelhető (és természetesen számítható is), hogy ebben az esetben a mért és számított adatok átlagos eltérése jóval nagyobb lenne, tehát a konkrét esetben a statisztika segít a döntésben. Másrészt, ha hat adatból kettőt el kell hagyni, akkor már az egész illesztés megkérdőjelezhető. Megbízható eredményhez legalább tíz, de inkább húsz illesztett adat szükséges, így a példánk csak az illesztés megvalósítási módjának az illusztrációja!

Az F3.2 táblázat az ábrához tartozó munkalap megfelelő részét mutatja be. Az eredeti adatsor az **A4..B9** tartományban található. Ezek közül a négy illesztett adatról készítettünk másolatot (az **A10..A13** tartomány a független, míg a **C10..C13** a függő változó értékeit tartalmazza).

Megjegyzés: Az illesztés előtt természetesen nem tudtuk, hogy mely adatokat kell elhagyni, ezért kezdetben az összes adatot átmásoltuk az **A10..A15**, valamint a **C10..C15** tartományokba. Az első illesztés alapján először elhagytuk az (1,0; 1,22) adatpárt úgy, hogy az eredeti adatsor **A12..C15** részét átmozgattuk (a MoveCell parancs segítségével) az **A11..C14** tartományba, és ezzel az elhagyandó adatpárt felülírtuk. Az illesztést újból végrehajtottuk, immár eggyel kevesebb adatra. A (2,5; 3,28) adatpár esetében hasonlóan jártunk el.

A végső illesztés eredményeit tartalmazza az **E1..F12** tartomány. Látható, hogy a tengelymetszet standard hibáját — amit a gyakorlaton megkövetelünk — a program nem számolja automatikusan, ezért az **F14** cellában megadtuk az S_b definícióját leíró képletet (ld. az F3.1 táblázatot). Ennek beírásakor három célszerű rövidítést használtunk fel:

1. A $\sum_{i=1}^{n} x_i^2$ kifejezést értékét a program már előzetesen kiszámolta és tárolta az **F4** cellában. A definícióban felhasználtuk ezen cella tartalmát.
2. Az illesztett adatok száma is megtalálható, mégpedig az **F7** cellában.
3. Az átlagos eltérés (jelölése σ az F3.1 táblázatban és **RegErr** az **AS-EASY-AS...** munkalapon), valamint a négyzetösszeg (S) F3.1 táblázatban megtalálható összefüggését használtuk fel S számítására, mivel a táblázatkezelő csak σ értékét adja meg. Természetesen használhattuk volna közvetlenül az S-et definiáló képletet is, de ez bonyolultabb függvénydefiníció beírását eredményezte volna.

A munkalapon az adatok elhelyezése nem véletlen, hanem az ábrakészítést könnyíti meg. Az F3.4 ábrán három adatsor van feltüntetve: az összes adatpár (x), az illesztésben felhasznált adatpárok (o) és a számított adatok egyenessel összekötve. Az **AS-EASY-AS...** minden ábrázolt adatsorhoz ugyanazokat a független változó adatokat rendeli, ezért a függő változók adatait úgy célszerű kijelölni és elrendezni, hogy a táblázatból is könnyen látható legyen, mit minek a függvényében ábrázolunk.

F3.2 táblázat: Egy tipikus munkalap egyenesillesztéshez.

	A	B	C	D	E	F
1	X	Yall	Yfit		Intercept	0.512307
2					Slope	0.991538
3	0.0				R^2	0.999799
4	0.5	1.02			Sum X^2	15.5
5	1.0	1.22			Sum Y^2	23.4009
6	1.5	1.99			Sum X*Y	18.955
7	2.0	2.48			Count	4
8	2.5	3.28			σx^2	0.8125
9	3.0	3.50			σy^2	0.798968
10	0.5		1.02		RegErr	0.017920
11	1.5		1.99		SlopeErr	0.009940
12	2.0		2.48		Formula:	†3.486923
13	3.0		3.50			
14					S_b	‡0.019568

†Az `F12` mező a következő definíciót tartalmazza: `0.512307692+0.991538462*@X`
‡Az `F14` cellában az alábbi definíció található: `@SQRT(F4*F10^2/(F7*F4-(@SUM(A10..A13))^2))`

Esetünkben az **`A3..A13`** tartományban[1] vannak a független változó adatai. Az első ábra függő változói a **`B3..B13`** mezőkön találhatók. Mivel a **`B3`** és **`B10..B13`** helyeken nincsen adat (csak üres mezők), itt a program nem rajzol semmit, csak a **`B4..B9`** adatok lesznek az **`A`** oszlop azonos sorban lévő pontjainak függvényében felrajzolva egy "x" szimbólummal. A második ábrázolandó pontsorozat tartománya **`C3..C13`**, így az előbbieknek megfelelően az illesztett adatok lesznek az ábrán egy "o" szimbólummal jelölve.

A harmadik adatsor, vagyis a számított egyenes megrajzolására a táblázatkezelő program egy speciális lehetőségét használtuk ki. Az illesztés végrehajtása után a program egy mezőbe (esetünkben az **`F12`**-be) teszi az egyenes egyenletét, a paraméterek kiszámított értékét felhasználva (ld. az **`F1`** és **`F2`** mezők értékét, összehasonlítva az **`F12`**-ben lévő definícióval). Ha ábraként csak ezt az egyenletet tartalmazó mezőt jelöljük ki, akkor a definíció alapján az **AS-EASY-AS...** kiszámítja a függvényértékeket a független változó tartományában (**`A3..A13`**) lévő minden pontra, és a számított adatpárokat felrajzolja. Azért írtunk az **`A3`** mezőbe egy nulla értéket, mert így a program a tengelymetszetnél is megrajzolja a számított egyenest.

[1] Emlékeztetőül: a független változók tartományát a \Graphics Range almenüben, az X gyorsbillentyűvel lehet definiálni, míg a lehetséges hat függő változó adatait ugyanebben az almenüben, az A, B, ..., F gyorsbillenyűkön keresztül.

II.F4. függelék
Hibaszámítás

A hibaterjedés számítása majdnem mindegyik gyakorlaton előfordul. A számítás menete, valamint a hiba/szórás értelmezése sokszor az egyik legnehezebb feladat a hallgatók számára. A szükséges elmélettel a [4,9] foglalkoznak megfelelő részletességgel, ebben a részben csak a hibák kiszámításához szükséges legfontosabb képleteket foglaltuk össze.

Sokak számára csak a négy alapműveletre (összeadás, kivonás, szorzás, osztás) alkalmazható leegyszerűsített szabály ismert: additív műveleteknél a szórások abszolút értékét, multiplikatív műveleteknél pedig a szórások relatív értékét kell összeadni, hogy az eredmény szórását megkapjuk. Ez az egyszerű eljárás jól alkalmazható a legegyszerűbb esetekben, de van két hibája: (a) mindig túlbecsüli az eredmény szórásának értékét és (b) a legfontosabb elemi függvényekre (pl. logaritmus, négyzetgyök) sem alkalmazható. A következőkben összefoglaljuk azokat a képleteket, amelyek segítségével a hibaszámítás korrekt módon elvégezhető.

Hibaterjedés számítása ismert szórású adatok egyszerű műveleteinél

Probléma

Legyen **k** darab független változónk, amelyek standard szórásukkal együtt megadott számok: $X_1 \pm S_{X1}, X_2 \pm S_{X2}, \ldots, X_k \pm S_{Xk}$. Ezekkel szeretnénk elvégezni tetszőleges — a $Z=f(X_1, X_2, \ldots, X_k)$ egyenlet által definiált — számításokat úgy, hogy Z-t és szórását (S_Z) is meg tudjuk határozni.

Megoldás

Tekintsük a független változókat normális eloszlású valószínűségi változóknak. Ekkor az eredményt (Z) — ami szintén egy valószínűségi változó — a matematika szabályai szerint kiszámítjuk a független változók értékeiből, ez lesz Z legvalószínűbb értéke. Z szórását a következő képlet alapján tudjuk számolni (a részletes levezetést és értelmezést ld. [9]-ben):

$$S_Z = \sqrt{\sum_{i=1}^{k} \left(\frac{\partial Z}{\partial X_i}\right)^2 S_{Xi}^2 + 2\sum_{j=2}^{k} \sum_{i=1}^{j-1} \frac{\partial Z}{\partial X_i} \frac{\partial Z}{\partial X_j} V_{X_i X_j}}\ , \qquad \text{(H1)}$$

ahol $V_{X_i X_j}$ az i-edik és j-edik független változó kovarianciája. Ezt az értéket kaphatjuk

a) számadatként programból (pl. egyenes illesztése);

b) tekinthetjük nullának, ha X_i és X_j valóban függetlenek és

c) végül számítható a következő képlet alapján, ha az X_i és X_j független változók értékei és szórásaik n darab adatból ($X_{i,1}$, $X_{i,2}$, ..., $X_{i,n}$) vannak kiszámítva:

$$V_{X_iX_j}=\frac{1}{n-1}\sum_{m=1}^{n}\left(X_{i,m}-\frac{1}{n}\sum_{q=1}^{n}X_{i,q}\right)\cdot\left(X_{j,m}-\frac{1}{n}\sum_{q=1}^{n}X_{j,q}\right) \qquad \text{(H2)}$$

Alapvető hibatranszformációk

Ha feltételezzük, hogy $X\pm S_X$ és $Y\pm S_Y$ egymástól függetlenek ($V_{XY}=0$), akkor az alapvető műveletekre a következő képletek vezethetők le (H1) felhasználásával (**a** és **b** tetszőleges, valós konstansok):

$$Z=aX \quad \Rightarrow S_Z=aS_X \quad \text{pl.} \quad 3(1{,}2\pm0{,}3)=3{,}6\pm0{,}9,$$

$$Z=aX\pm bY \quad \Rightarrow S_Z=\sqrt{(aS_X)^2+(bS_Y)^2} \quad \text{pl.} \quad \frac{1}{2}(2{,}2\pm0{,}2)-\frac{1}{4}(8{,}4\pm1{,}0)=-1{,}0\pm0{,}2(7),$$

$$Z=\pm aXY \quad \Rightarrow S_Z=\sqrt{a^2(Y^2S_X^2+X^2S_Y^2)} \quad \text{pl.} \quad 0{,}5(2{,}2\pm0{,}2)(8{,}4\pm1{,}0)=9{,}2(4)\pm1{,}3(8),$$

$$Z=\pm a\frac{X}{Y} \quad \Rightarrow S_Z=\sqrt{a^2\left(\frac{S_X^2}{Y^2}+\frac{X^2S_Y^2}{Y^4}\right)} \quad \text{pl.} \quad 2{,}0\frac{(2{,}2\pm0{,}2)}{(8{,}4\pm1{,}0)}=0{,}52\pm0{,}07(8),$$

$$Z=\pm\frac{a}{X} \quad \Rightarrow S_Z=\sqrt{a^2\frac{S_X^2}{X^4}} \quad \text{pl.} \quad \frac{2{,}0}{(8{,}4\pm1{,}0)}=0{,}24\pm0{,}02(8).$$

Az elemi és transzcendens függvények használatakor érvényes hibaterjedés is könnyen számolható. Néhány fontosabb, a gyakorlatokon is használható példa:

$$Z=\pm aX^b \quad \Rightarrow S_Z=\sqrt{(abS_XX^{b-1})^2} \quad \text{pl.} \quad -2{,}0(3{,}0\pm0{,}5)^{1{,}2}=-7{,}4(7)\pm1{,}4(9),$$

$$Z=\pm ae^{bX} \quad \Rightarrow S_Z=\sqrt{(abS_Xe^{bX})^2} \quad \text{pl.} \quad -2{,}0e^{1{,}2(3{,}0\pm0{,}5)}=-73{,}(2)\pm43{,}(9),$$

$$Z=\pm a\ln(bX) \quad \Rightarrow S_Z=\sqrt{a^2\frac{S_X^2}{X^2}} \quad \text{pl.} \quad -2{,}0\ln(1{,}2(3{,}0\pm0{,}5))=-2{,}56\pm0{,}33.$$

Természetesen a (H1) és esetlegesen a (H2) képlet alkalmazásával tetszőleges esetben le tudjuk vezetni a szükséges képleteket.

Felhasznált és ajánlott irodalom

1. Fizikai-kémiai laboratóriumi gyakorlatok, szerk. Peintler Gábor
   JATEPress, Szeged, 1998
2. P. W. Atkins: Fizikai kémia, I–III. kötet,
   Tankönyvkiadó, Budapest, 1992.
3. Erdey-Grúz Tibor, Schay Géza: Elméleti fizikai kémia, I–III. kötet,
   Tankönyvkiadó, Budapest, 1962.
4. Erdey-Grúz Tibor, Schay Géza: Fizikai-kémiai praktikum, I–II. kötet,
   Tankönyvkiadó, Budapest, 1965.
5. Bevezetés a fizikai kémiai mérésekbe, I–II. kötet, szerk. Kaposi Olivér
   Tankönyvkiadó, Budapest, 1988.
6. Németh Béla: Kémiai táblázatok,
   Műszaki Könyvkiadó, Budapest, 1979.
7. Dobos Dezső: Elektrokémiai táblázatok,
   Műszaki Könyvkiadó, Budapest, 1984.
8. Analitikai zsebkönyv, szerk. Mázor László
   Műszaki Könyvkiadó, Budapest, 1978.
9. A. C. Norris: Computational Chemistry,
   John Wiley & Sons, New York, 1981.
10. G. Gran, *Analyst*, 77. köt., 661. old., **1952**
11. H. M. Irving, *Analytica Chimica Acta*, 38. köt., 475. old., **1967**