

Függelék

általános kémia laboratóriumi gyakorlatokhoz

Összeállította: Dr. Schusztér Gábor

Tartalomjegyzék

1. Műszaki-tudományos ábrák készítése	2
2. Egyenes illesztése és statisztikai jellemzése MS Excel-ben	5
2.1. Alapfogalmak	5
2.1.1. Egyenesillesztés elve	5
2.1.2. Excel trendvonal	5
2.1.3. Lineáris regresszió / paraméterbecslés MS Excelben	7
2.2. Feladatok megoldásának részlépései	8

1. Műszaki-tudományos ábrák készítése

Mind a kézzel, milliméterpapíron elkészített, mind a számítógépes ábrákkal szemben ugyanazok a követelmények:

- Lehetőleg minden mért adat, vagy azok transzformáltjai szerepeljenek az ábrán.
- Legyen megfelelő – szükség esetén mértékegységgel ellátott – címe mind a tengelyeknek, mind az ábrának. A feliratok mind szakmai, mind nyelvtani szempontból legyenek helyesek. Lehetőleg név és dátum is szerepeljen az ábrán.
- A tengelyek beosztásának és a címkefeliratoknak olyanoknak kell lenniük, hogy az adatok könnyen ábrázolhatók és visszaolvashatók legyenek, és minimalizálják az ábra haszontalan területeit. 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 és/vagy adatsor együttes ábrázolása esetén az egyes görbék vagy adatsorok legyenek világosan elkülöníthetők.

Természetesen lehetnek egyéb elvárások is a konkrét feladattól függően. Ritkán a követelmények egyike-másika nem teljesíthető 100 %-osan (pl. ha egy rossz pont nagyságrendekkel különbözik a többitől, az teljesen eltorzítja az ábrát), de az esetek túlnyomó többségében a fentiek betartása elegendő hibátlan ábrák készítéséhez.

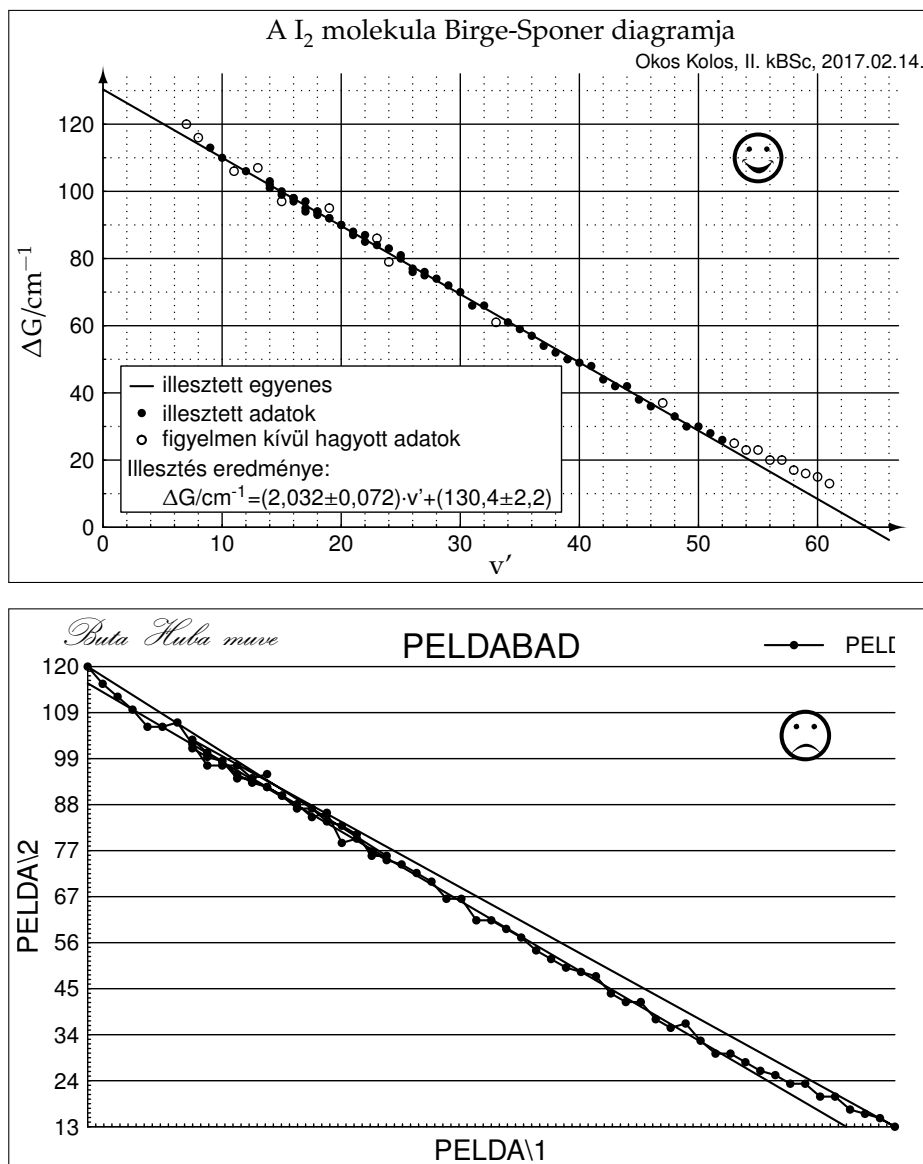
Az eddigiekből is kitűnhet, hogy az ábrakészítő programok felületes ismerete sokszor nem elég. *Általában nem lehet elfogadni kész ábraként, amit egy program az adatok bevitel után az alapbeállításával felrajzol a képernyőre, hanem olyan szinten kell megtanulni az alkalmazott program kezelését, hogy a fenti követelmények teljesíthetők legyenek!* A tudományos életben ezt fokozottan kell hangsúlyozni, mert a legtöbb kereskedelembe 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 alábbiak egy példán keresztül mutatják be azokat a jellemző hibákat, amiket a leggyakrabban szoktak elkövetni számítógéppel történő ábrakészítés során. Az 1. ábra alsó és felső panelje ugyanarra az adatsorra történő egyenesillesztést illusztrálja. A felső panel teljes mértékben megfelel a fentebb részletezett követelményeknek, míg az alsó panel a – tapasztalat szerint – leggyakrabban előforduló hibákat mutatja. Ezek könnyen elkerülhetők a használt program megfelelő szintű ismeretével. A függelék ezen szakaszának további része a két panel 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, főleg közgazdasági grafikonokon. Tudományos ábrákon vonallal illesztett görbét szokás jelölni, így a mért pontok összekötése félrevezető. Ráadásul érthetetlen ábrákhoz vezethet, ha az adatok nincsenek szigorúan növekvő vagy csökkenő sorrendben, pl. az 1. ábra alsó paneljében egyetlen nem sorrendben lévő pont egy felesleges vonalat ad.

Rossz tengelytartomány: Néhány program automatikusan ráteszi az ábrára a koordináta-rendszer origóját. Az ábrázolandó pontok tartományától függően ez ahhoz vezethet, hogy az ábra kicsiny részére zsúfolódik össze minden pont, azok menete kivehetetlen lesz.

Egyetlen tengelybeosztás: Sok program a tengelyek minimális és maximális értékét az adatsorokból kapható minimális és maximális értékhez rendeli. Az alsó panelen az y-tengely beosztása rossz, mert a 13 – 120 tartományt nem lehet jól felosztani tíz részre. Ráadásul a beosztások feliratai pontatlanok,



1. ábra. Egy kifogástalanul (felső panel) és egy a tipikus hibákat bemutató (alsó panel) számítógépes ábra.

csak egész értékre vannak megadva. Emiatt hibás a visszaolvasás, azonos hosszúságú tartományokhoz eltérő értékek tartoznak (pl. $120 - 109 \neq 109 - 99$)! A felhasználónak tudnia kell, hogyan lehet beállítani a tengelyek minimális és maximális értékeit, a beosztás sűrűségét és a beosztások helyét jelző számok kiírati formátumát!

Automatikus tengelyválasztás: Emiatt rossz beosztás, értelmetlen vagy hiányzó címek és beosztásfeliratok lehetnek a tengelyeken. Az alsó panelen hiányoznak az x -tengely beosztásainak feliratai, a semmitmondó automatikus tengelycímek az adatállomány nevéből és a felhasznált oszlopok sorszámából adódnak, a szélső adatok pedig szinte „lelőgnak” 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: Bosszantó információvesztés lehet. Példánkban a rossz ábrán a dátum hiányzik.

Rossz pozícionálás: Az ábra valamely részén szerencsésebb esetben csak komikus, rosszabb esetben információvesztéshez vezethet. Esetünkben a jelmagyarázó blokk annyira rossz helyen van az alsó panelen, hogy a fele lemaradt.

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

Rácsozat: Ha van, gondosan kell beállítani. Nem segíti az ábra olvasását a túl sűrű rácsozat, mert szinte elfedi a görbéket. A ritka rácsozat sem jó, mivel a függvényértékeket nagyon nehéz visszaolvasni ilyen esetekben. Sokszor tisztább az ábra, ha egyáltalán nincs rácsozat. Az biztosan nem jó, ha vagy csak a vízszintes, vagy csak a függőleges rácsvonalak vannak feltüntetve, ahogy azt a rossz ábra mutatja.

Nem megfelelő betűtípus és / vagy betűnagyság használata: Jobb esetekben csak csúnya vagy komikus feliratokhoz vezet, rosszabb esetben félreértésre is alkalmat adhat. Az 1. ábra alsó paneljén a név értelmetlenül csicsás és nem ékezetes betűkkel készült. Ábrákon általában érdemes egyszerű vonalvezetésű és vastagabb betűtípusokat használni (pl. Swiss, Arial, Helvetica, Tahoma, Verdana, Calibri, stb. típusok).

Az elhagyott pontok nem szerepelnek az ábrán, vagy jelöletlenek: Ha nem tüntetjük fel az ábrán az illesztésnél figyelmen kívül hagyott pontokat, akkor információt veszünk a mérések valódi pontosságáról és a pontelhagyások okairól. Ha a rossz pontokat az illesztettekkel azonos módon jelöljük (ahogy ez a rossz ábrán látható), akkor a számított adatok reprodukálása lesz nehéz.

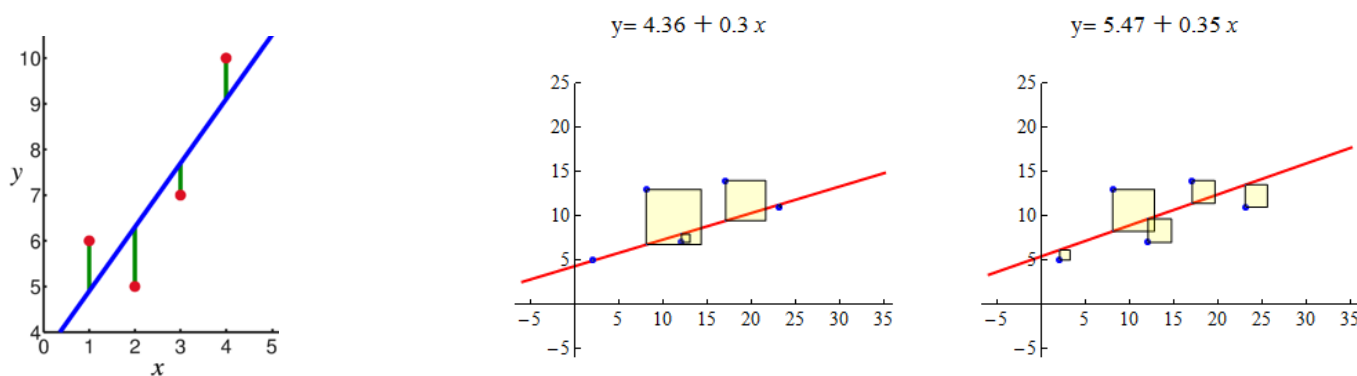
2. Egyenes illesztése és statisztikai jellemzése MS Excel-ben

2.1. Alapfogalmak

2.1.1. Egyenesillesztés elve

- Az "egyenesillesztés" kifejezés alatt lineáris regressziót, vagy más szóval lineáris paraméterbecslést értünk (lásd 2. ábra, bal oldal).
- Az illesztendő egyenes egyenlete kétféle lehet: 1) $y = a \cdot x$ (azaz az y-tengellyel való metszete az origó) és 2) $y = a \cdot x + b$ (azaz az y-tengellyel való metszete illesztendő paraméter). A mérés elvéből következik, hogy melyiket kell választani. Ha pl. a Lambert–Beer-törvényt tekintjük, azaz $A = \varepsilon c \ell$, akkor az $A - c$ függvényre $x = 0$ -ban $y = 0$ (hiszen ha nincs jelen az elnyelő anyag, akkor elnyelése sem lehet).
- A legjobb illeszkedés eléréséhez a programok gyakran a *legkisebb négyzetek módszerét* alkalmazzák. Ekkor az eltérés négyzetösszeg minimalizálása történik (lásd 2. ábra, jobb oldal). Egyéb függvények (pl. polinom, exponenciális stb.) is alkalmazható az eljárás.

Általános Kémia laborgyakorlaton mm-papíron történő illesztéskor: Úgy húzzák be az egyenest, hogy az összes mérési pont között, azokhoz a lehető legközelebb haladjon el!



2. ábra. **Bal oldal:** Lineáris regresszió: mért értékek (piros kör), feltételezett függvénykapcsolat (kék egyenes), és a köztük levő véletlen eloszlású eltérés (zöld összekötő). **Jobb oldal:** Lineáris paraméterbecslés legkisebb négyzetek módszerével; rossz (bal) és jó (jobb) illesztés.

2.1.2. Excel trendvonal

- Csak illusztrációra alkalmazandó!* Megadott függvénykapcsolat alapján a mérési adatokból ábrát kell készíteni (minden adat ábrázolandó, de nem minden adat illesztendő!), majd erre kell felvenni a lineáris trendvonalat.
 - Paraméterbecslést megelőző ábrakészítés *haszna*
 - Függő és független változó helyes azonosításának ellenőrzése (az adatpárok olyan trendet követnek, amit *józan megítélés alapján* várunk)
 - Adatok között potenciálisan fennálló lineáris függvénykapcsolat vizuális mérlegelése
 - Felismert trendbe nem illeszkedő mérési adatok azonosítása
 - Függő & független változók azonosítása: Pl.: Ha egy A anyag $A \rightarrow B + C$ séma szerint elsőrendű reakcióban bomlik, akkor az integrált sebességi egyenlet $[A]_t = [A]_0 \cdot e^{-kt}$ alakú. A mért $[A]_t - t$ adatpárokból $[A]_0$ és k illesztendő paraméterek többféleképpen is meghatározhatók. Amennyiben

lineáris paraméterbecslés a cél, vegyük az egyenlet mindkét oldalának természetes alapú logaritmusát

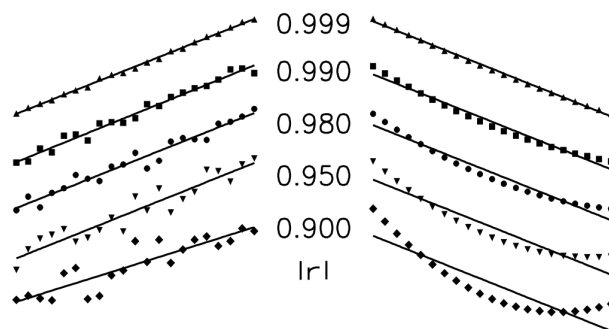
$$\ln[A]_t = \ln[A]_0 - kt \quad (1)$$

Ezt követően (minden egyéb illesztendő egyenlet esetén is) az a feladat, hogy az egyenes

$$y = a \cdot x + b \quad (2)$$

általános egyenletének megfelelően azonosítsuk az egyes paramétereket és változókat.

- * A valós mérés során t halad előre (ha tetszik, mi hagyjuk telni, azaz önkényesen tudjuk változtatni) és ennek következtében $[A]_t$ változik, azaz az (1) egyenletben t a független és $\ln[A]_t$ a függő változó.
 - * A (2) egyenletben a független változónak (x) szorzótényezője van (a), mely az (1) egyenlet szerint $-k$ (negatív előjel!).
 - * Végül a (2) egyenlet szerinti additív tag (b) megfelelője az (1) egyenletben $\ln[A]_0$.
 - * Ezen logikát követve bármely (linearizálható) egyenlet esetén azonosíthatjuk a változókat és az illesztendő paramétereket még akkor is, ha azok összetettebb matematikai kifejezések.
- A trendvonal ábrára történő felvételével megtörténik az adatsorra legjobban illeszkedő egyenes kirajzolása és az illesztett paraméterek kiírására (meredekség, tengelymetszet, determinációs együttható). A probléma az, hogy az illesztett egyenes statisztikai adatait nem kapjuk meg, így a továbbiakban nem lesz módunk pl. a szórások számítására.
 - A trendvonal felvételekor egy további opció: egyparaméteres illesztés, azaz a tengelymetszet értékének önkényes rögzítése (0 vagy attól eltérő).
 - Korrelációs (r) & determinációs együttható (R^2) értelmezése
 - Korrelációs együttható (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 (y) és független (x) változók adatai között. Minél közelebb van $|r|$ értéke egyhez, annál valószínűbb a lineáris kapcsolat. *Ezt a paramétert a laboratóriumi gyakorlatokon nem számoljuk, feltételezzük, hogy az elméleti lineáris kapcsolat valóban fennáll.*
 - Determinációs együttható (R^2): Az R^2 számítási módja és a kapott érték értelmezése az illesztett függvény fajtájától (egyenes, exponenciális stb.) és a független változók számától is függ. A laboratóriumi gyakorlatokon egyváltozós (x), és egy- (meredekség) vagy kétparaméteres (meredekség, tengelymetszet) egyenesillesztést hajtunk végre. Ekkor $R^2 = r^2$. A determinációs együttható *nem* azt mutatja meg, hogy az egyenes mennyire jól írja le az adatsorban észlelt trendet, hanem azt számszerűsíti, hogy az illesztett egyenes körül *milyen mértékben szórnak* a mért adatok. Amennyiben az adatsor valóban lineáris trendet követ, az 1-hez minél közelebbi R^2 egyre kisebb szórással terhelt mérést tükröz.
 - Az adatsor illesztéshez használt függvény megfelelősége, és az adatsor linearitásának mértéke közötti különbséget szemlélteti a 3. ábra. Látható, hogy egy adott R^2 értékhez két teljesen különböző illesztés is rendelkezhető. A bal oldali pontseregek valóban lineáris trendet követnek így R^2 értékének 1-hez közeli volta az adatsor kis szórását mutatja. Ezzel szemben a jobb oldali pontseregekre történő egyenesillesztés teljesen helytelen, noha ugyanazon R^2 értékeket eredményezi. *Végeredményben R^2 csak akkor jelent az egyenes illeszkedésére, vagy az adatsor szórására nézve hasznos információt, ha az illesztett trend valóban lineáris.* Az R^2 1-hez közeli értéke önmagában nem jelenti, hogy az adatsort lineáris trend jellemzi.



3. ábra. A korrelációs együttható abszolút értékének ($|r|$) szemléltetése. **Bal oldal:** normális eloszlású hibával terhelt adatok; **jobb oldal:** szisztematikus görbülettel rendelkező adatsorok. Ezen illesztések esetén $R^2 = r^2$, az ábra értelmezésekor nem jelent változást, hogy $|r|$ és nem R^2 került feltüntetésre.

2.1.3. Lineáris regresszió / paraméterbecslés MS Excelben

- Amennyiben nem elegendő a trendvonal alkalmazása, mert az illesztett egyenes statisztikai adataira is szükségünk van, Excel-ben a *LIN.ILL* függvényt használjuk.
- Ez egy tömbfüggvény, azaz a bemeneti és a válaszerőtekek több cellára is kiterjednek. A függvény használatakor CTRL+SHIFT+ENTER billentyőkombinációra van szükség.
A laborgyakorlatokon egyváltozós (x), és egy- (meredekség) vagy kétparaméteres (meredekség, tengelymetszet) egyenesillesztéskor 2 (oszlop) $\times$ 5 (sor) cellából álló tömböt kell előre kijelölni és a következő parancsot kell begépelni:

$$= \text{LIN.ILL}(\text{ismert_y}; \text{ismert_x}; \text{konstans}; \text{stat}),$$

majd CTRL+SHIFT+ENTER billentyőkombináció.

- $\text{ismert_y}; \text{ismert_x}$ $\rightarrow$ A függő (y) és a független (x) változókat tartalmazó cellák kijelölése. A kijelölésben nem lehet megszakítás, ezért rossz mérési adatokat is tartalmazó adatsor esetén (ezt az előzetes ábrakészítés alapján ítéltethjük meg) az illesztendő adatpárokat új cellákba kell másolni.
- konstans $\rightarrow$ Ha értéke 1 / IGAZ: a paraméterbecslés során a tengelymetszet (b) értéke meghatározásra kerül; ha értéke 0 / HAMIS: a paraméterbecslés során a tengelymetszet (b) értéke 0 és az illesztés $y = a \cdot x$ egyenlet alapján történik (a mérés elve a meghatározó, pl. oldat abszorbanciája $c = 0$ oldott anyag koncentrációnál).
- stat $\rightarrow$ célszerű értéke 1 (különben a függvény nem számolja a kívánt statisztikai adatokat).
- LIN.ILL*-el meghatározott statisztikai adatok a tömbfüggvény kiíratási formátumának megfelelően (azaz az eredetileg kijelölt 2 (oszlop) $\times$ 5 (sor) cellában az alábbi értékek jelennek meg):

	1	2
1	meredekség	tengelymetszet
2	meredekség standard hibája	tengelymetszet standard hibája
3	R^2	<i>a gyakorlaton nem releváns</i>
4	<i>a gyakorlaton nem releváns</i>	szabadsági fokok száma
5	<i>a gyakorlaton nem releváns</i>	<i>a gyakorlaton nem releváns</i>

- Szabadsági fokok száma (S_z): egymástól független mérési adatok esetén $S_z = N - P$, ahol N az adatok száma és P a meghatározandó paraméterek száma.

– Standard hiba ($\neq$ szórás!), kapcsolatuk: szórás (σ) = $\sqrt{\text{szabadsági fok} \cdot \text{standard hiba}}$

- *Megjegyzés:* Az MS Office 365 Excel verziótól a *LIN.ILL* (és a többi tömbfüggvény is) megadható lineáris egyenletként, azaz nem kell előre kijelölni egy tömböt. Az Excel automatikusan lefoglalja a megfelelő számú cellát, s nem kell a CTRL+SHIFT+ENTER billentyűkombináció sem, elég az ENTER. A kompatibilitás miatt a fentebb leírt *régi* módszer továbbra is működik.

2.2. Feladatok megoldásának részlépései

1. A függő (y) és a független (x) változók azonosítása a linearizált egyenlet és a szöveg alapján.
2. Az illesztendő paraméter(ek) (meredekség, tengelymetszet) azonosítása a linearizált egyenlet és a szöveg alapján.
3. Amennyiben a mérési adatok nem közvetlenül egyeznek meg az azonosított függő (y) és független (x) változókkal, úgy ezeket külön oszlopokban ki kell számítani a rendelkezésre álló mérési adatokból.
4. Amikor a függő (y) és független (x) változókat tartalmazó adatsorok (oszlopok) már rendelkezésre állnak, egy ábrát kell készíteni a linearizált egyenletnek megfelelően. Az ábra ekkor tartalmazza az összes mérési adat alapján meghatározott ábrázolandó adatpárt. Az ábra alapján azonosítani kell a trendet nem követő, *kiugró* adatpárokat, melyeket az illesztéshez nem szabad figyelembe venni.
5. Az illesztendő (*kiugró adatokat* már nem tartalmazó) adatpárokat célszerű új oszlopokba is átmásolni, majd az előbbieken készített ábrán feltüntetni (a *kiugró adatokat* is tartalmazó adatsor mellett). Az ábrán meg kell adni, hogy melyik jelölő melyik adatsorhoz tartozik (jelmagyarázat). Az ábrán a trendvonalat az illesztendő adatokat tartalmazó adatsorra kell felvenni.
6. Az illesztendő adatokat tartalmazó oszlopok segítségével lineáris paraméterbecslést kell végrehajtani a *LIN.ILL* Excel-függvénnyel. Az elkészített ábrán kapott trendvonal és a *LIN.ILL*-ből kapott illesztési paraméterek összehasonlítása segíti annak megítélését, hogy mindent megfelelően hajtottunk-e végre.
7. Ellenőrizni kell, hogy mind az adatokat tartalmazó oszlopok, mind a készített ábra tudományos minőségű, érthető, és tartalmazza a szükséges magyarázatokat, mértékegységeket.
8. A *LIN.ILL*-ből kapott standard hibák alapján ki kell számítani a szórásokat (tengelymetszet, meredekség).
9. A linearizált egyenletnek és a szövegnek megfelelően további számítások elvégzésére lehet szükség, hogy az illesztett paraméterektől eljussunk a meghatározni kívánt mennyiségekig. Ügyeljünk arra, hogy potenciálisan mind az illesztések, mind a számítások során dimenzióval rendelkező mennyiségekkel dolgozunk, mely meghatározza a végeredmény mértékegységét is.
10. Ha a feladat kéri, figyelembe kell venni a megfelelő szórásterjedéssel kapcsolatos összefüggéseket is.