Tractament de Dades

Avaluació d'incerteses. Representació Gràfica de Dades. Regressions lineals.

Index

- Concepte d'error en una mesura.
- Tractament d'errors:
 - Errors instrumentals
 - Errors aleatoris
 - Magnituds derivades: propagació d'errors
- Representació correcta de dades en una gràfica.
- Aplicació del mètode dels mínims quadrats per a ajustar dades experimentals a una recta.

Concepte d'error en una mesura

Mesures experimentals En les mesures experimentals sempre es produeixen errors o incerteses. La natura d'aquestes incerteses és d'origen divers, i va des del fet que els nostres aparells tenen una precisió finita (incertesa instrumental), fins a variacions en l'ambient on es realitza la mesura —com per exemple canvis en la temperatura— (errors aleatoris). Evidentment també existeixen els errors pròpiament dits que es produeixen quan fem malament la mesura. Aquests darrers s'eviten posant una cura extrema quan es fan els muntatges experimentals. Els altres són inevitables i és necessari donar-los un tractament adequat. En aquest guió parlarem sempre d'errors quan ens referim a errors o incerteses experimentals.

Les mesures experimentals sempre han d'anar acompanyades del seu respectiu error, per exemple si dos grups experimentals ${\bf A}$ i ${\bf B}$ mesuren una mateixa resistència i donen com a resultat

Grup **A**:
$$R_A = 1215 \pm 1 \Omega$$
,
Grup **B**: $R_B = 1230 \pm 20 \Omega$,

veiem que els dos grups estan d'acord ja que el valor del grup \mathbf{A} (R_A) es troba entre el valor mínim ($R_{B\,\text{min}} = 1210~\Omega$) i el valor màxim ($R_{B\,\text{max}} = 1250~\Omega$) del grup \mathbf{B} . Per altra banda mirant aquestes mesures és clar que el valor del grup \mathbf{A} és molt més precís que el del grup \mathbf{B} . Si no tinguéssim els errors seríem incapaços de saber quin dels dos grups ha fet l'experiment millor, i, encara pitjor, no sabríem si els seus resultats estan d'acord o

no. Els errors es podem presentar com acabem de fer (en termes dels *errors absoluts*), o també es poden expressar en percentatge respecte la magnitud mesurada, i.e.

$$R_A = 1215 \Omega \pm 0.08 \%$$
, $R_B = 1230 \Omega \pm 1.6 \%$.

En aquest cas parlem d'errors relatius de les magnituds mesurades.

0.0.1 Tractament d'errors

Errors instrumentals. Aquests tipus d'errors estan associats a la precisió finita que tenen els aparells. Si tenim un regle amb una escala que arriba fins a 1 milímetre és clar que la màxima precisió amb què podem mesurar amb aquest regle és de 1 mm. És a dir, si amb aquest regle mesurem una distància de 10.5 cm. la mesura experimental es donarà com:

$$L = 10.5 \pm 0.1 \,\mathrm{cm}$$
.

En general,

L'error instrumental (ε_i) vindrà donat per l'escala mínima que és capaç de mesurar l'aparell.

Errors aleatoris: tractament estadístic. Com hem dit en l'apartat de mesures experimentals poden existir causes ambientals que modifiquin el valor de la magnitud a mesurar. Per exemple, en les mesures de longitud, un canvi de temperatura pot fer dilatar tant l'objecte que volem mesurar com el propi regle. Aquests errors no es poden evitar. La manera de tractar-los és realitzar diverses mesures de la mateixa magnitud i fer un tractament estadístic dels valor obtinguts.

Veiem primer un exemple. Suposem que fem 3 mesures d'una longitud i obtenim els següents resultats:

$$L_1 = 10.5 \,\mathrm{cm}$$
, $L_2 = 10.3 \,\mathrm{cm}$, $L_2 = 10.7 \,\mathrm{cm}$,

cadascuna amb un error instrumental $\varepsilon_i = 0.1 \, \mathrm{cm}$. El valor mesurat de la longitud és la mitjana aritmètica de les mesures individuals:

$$\langle L \rangle = \frac{10.5 + 10.3 + 10.7}{3} = 10.5 \text{ cm} .$$

L'error estadístic o desviació típica s'obté a partir de les distàncies entre la mitjana aritmètica i els diferents valors individuals. La desviació típica indica com de disperses al voltant de la mitjana estan les nostres mesures (si tenim totes les mesures molt concentrades al voltant de la mitjana tindrem poca desviació típica; de manera semblant, si tenim mesures molt disperses o allunyades entre sí, tindrem una desviació típica gran). Utilitzarem la millor estimació de la desviació típica (σ) . Concretament el quadrat de la millor estimació de la desviació típica en aquest cas és²:

¹En els diferents llibres s'utilitzen diferents criteris sobre com calcular aquesta magnitud, nosaltres utilitzarem la donada en l'expressió (2).

²Estrictament aquest tractament només es pot aplicar quan tenim més de 5 mesures d'una mateixa magnitud.

del qual s'obté

$$\sigma = \sqrt{\sigma^2} = 0.11547005384 \text{ cm}$$
.

El sentit de la desviació típica és que el *veritable* valor de la magnitud es troba entre $\langle L \rangle - \sigma$ i $\langle L \rangle + \sigma$ amb una probabilitat del 68%³. Per trobar l'error total (ε_T) en la mesura cal *combinar* l'error instrumental $(\varepsilon_i = 0.1 \, \text{cm})$ i l'error estadístic:

$$\varepsilon_T^2 = \varepsilon_i^2 + \sigma^2 = 0.02333 \text{ cm}^2 \implies \varepsilon_T = 0.1527... \text{ cm}$$

ara bé: no té sentit donar un error amb tantes xifres significatives, per tant l'aproximarem sempre a una o dues xifres significatives. Finalment el resultat experimental de la mesura seria (agafant dues xifres significatives):

$$L = 10.50 \pm 0.15 \,\mathrm{cm}$$
.

Expressem ara tot el procés de manera abstracta:

- Fem N mesures d'una mateixa magnitud $X: X_1, X_2, \ldots, X_N$. Cadascuna té un error instrumental ε_i .
- El valor mesurat de la magnitud és la mitjana aritmètica de les mesures individuals:

$$\langle X \rangle = \frac{1}{N} \sum_{j=1}^{N} X_j \ . \tag{1}$$

• La desviació típica representa l'error estadístic de la mostra de mesures. Podem estimar la desviació típica fent servir un *estimador*. El millor estimador de la desviació típica és

$$\sigma = \sqrt{\frac{1}{N(N-1)} \sum_{j=1}^{N} (X_j - \langle X \rangle)^2} = \sqrt{\frac{1}{N-1} (\langle X^2 \rangle - \langle X \rangle^2)}$$
 (2)

on $\langle X^2 \rangle$ es la mitja dels valors de $X_j^2,$ calculat com en l'expressió (1) substituint X_j per $X_j^2.$

• L'error total és la combinació de l'error instrumental amb l'estadístic:

$$\varepsilon_T = \sqrt{\varepsilon_i^2 + \sigma^2} \tag{3}$$

• El valor de la mesura és:

$$X = \langle X \rangle \pm \varepsilon_T \ . \tag{4}$$

 $^{^3}$ La probabilitat que la magnitud es trobi en un interval $\langle L \rangle \pm 2\sigma$ serà del 95%. La probabilitat que la magnitud caigui en l'interval $\langle L \rangle \pm 3\sigma$ serà del 99%, etc.

Errors en magnituds derivades: propagació d'errors. Tot el que hem vist als apartats anteriors ens permet fer el tractament d'errors sobre mesures directes, però en la majoria de situacions el que es fa són *mesures indirectes*, per exemple per mesurar una resistència en un circuit el que fem és mesurar el voltatge i la intensitat, i a partir de la llei d'Ohm

 $R = \frac{V}{I}$

calculem la resistència. El problema és: com trobem l'error associat a R coneixent els errors en V i en I? La resposta és que hem de propagar els errors. Donarem primer l'expressió general per la propagació d'errors i en veurem després un exemple.

• Per calcular l'error en la determinació d'una magnitud derivada F que depèn de diverses magnituds x_i (j = 1 ... n) segons una funció

$$F = f(x_1, \dots, x_n) ,$$

i tenint en compte que cada magnitud x_j té un error diferent ε_j utilitzarem la següent expressió:

$$\varepsilon_F = \sqrt{\sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \varepsilon_j\right)^2} , \qquad (5)$$

on $\partial f/\partial x_j$ és la derivada de la funció f respecte la variable x_j mantenint les altres variables constants.

Veiem com s'aplica això en el cas de la mesura de la resistència. Suposem que tenim un circuit amb una resistència i mesurem el voltatge i la intensitat de corrent que passen pel circuit:

$$V = 1.5 \pm 0.1 \,\text{V}$$
, $I = 3.0 \pm 0.3 \,\text{mA} = (3.0 \pm 0.3) \times 10^{-3} \,\text{A}$.

En aquest cas les variables x_j són V i I, i la funció f és

$$R = f(V, I) = \frac{V}{I} .$$

El valor de la resistència és

$$R = \frac{1.5 \text{ V}}{3.0 \text{ mA}} = \frac{1.5 \text{ V}}{3.0 \times 10^{-3} \text{A}} = 500 \text{ }\Omega \text{ }.$$

Les diferents derivades que apareixen en l'expressió (9) són:

$$\frac{\partial f}{\partial V} = \frac{1}{I} , \quad \frac{\partial f}{\partial I} = -\frac{V}{I^2} ,$$

i per tant el quadrat de l'error és

$$\varepsilon_R^2 = \left(\frac{1}{I}\,\varepsilon_V\right)^2 + \left(-\frac{V}{I^2}\varepsilon_I\right)^2$$

$$= \left(\frac{1}{3.0 \times 10^{-3} \text{A}} \, 0.1 \text{V}\right)^2 + \left(-\frac{1.5 \text{V}}{(3.0 \times 10^{-3} \text{A})^2} \, 0.3 \times 10^{-3} \text{A}\right)^2$$

$$= 3611.11 \,\Omega^2$$

i fent-ne l'arrel quadrada obtenim

$$\varepsilon_R = 60.0925$$
.

Seguint el criteri d'arrodonir l'error a les 2 primeres xifres significatives, tenim finalment que el valor mesurat de la resistència és:

$$R = 500 \pm 60 \Omega$$
.

Vegeu l'apèndix A, on trobareu les fórmules de propagació d'errors que haureu d'utilitzar en les pràctiques.

Representació gràfica de dades

Les dades experimentals s'han de presentar, sempre que sigui possible, en forma de taules, recordant que en una mateixa columna s'han de posar diferents mesures d'una mateixa magnitud. Per altra banda, recordeu que

El resultat d'una mesura **sempre** s'expressa com un valor, la incertesa associada i la unitat.

En molts casos resulta molt útil representar les dades en forma de gràfica, ja que d'aquesta manera podem, a primer cop d'ull, fer-nos una idea de quina és la relació entre les variables involucrades. És més, mitjançant una gràfica podem determinar la relació empírica entre les variables (és a dir, podem trobar una "fòrmula" que relaciona les variables).

En fer una representació gràfica d'un conjunt de dades cal tenir en compte les següents consideracions:

- S'ha de seleccionar el tipus de paper adequat. El més frequent és utilitzar paper millimetrat, però hi ha situacions en què és més pràctic utilitzar paper semilogarítmic o bé paper logarítmic.
- S'han de seleccionar i denominar les escales coordenades:
 - Seleccioneu els eixos: la variable independent en l'eix de les x (abscisses) i la dependent en l'eix de les y (ordenades).
 - Establiu una escala adequada per a una lectura fàcil: 1 quadre = $1, 2, 5, 10, \ldots$ unitats. Eviteu relacions de l'estil: 1 quadre = $3, 6, 7, \ldots$ unitats.
 - Deixeu un cert marge entre els eixos i el límit del paper.
 - Ajusteu correctament l'escala per tal de no posar les dades en una part petita de la gràfica.
 - En el cas de representar una recta, el pendent geomètric s'ha d'acostar a la unitat (i.e. la recta ha de fer un angle d'aproximadament 45°). Recordeu que el pendent geomètric (el de la recta que dibuixeu) i el real (el quocient entre la variable y i la x) no tenen per què ser iguals.
 - Etiqueteu els eixos amb les magnituds i unitats corresponents.
 - Marqueu només cada 2, 4 o 5 quadres.

- Els punts que representen els dades experimentals s'han de poder localitzar i visualitzar fàcilment, indicant les barres o rectangles d'incertesa. Vegeu la figura 1 com a exemple.
- En el cas d'ajustar els punt experimentals a una corba, hem d'unir els punts experimentals amb línies suaus, *mai* fent ziga-zagues.
- La gràfica ha d'estar correctament titulada: s'ha de posar un títol de la forma "Y en funció de X", substituint Y i X pel que correspongui. En el cas d'haver ajustat les dades experimentals a una funció, s'ha d'indicar el tipus de funció (posant la seva equació), els valors i incerteses dels paràmetres de la funció, i el valor del coeficient de correlació r.

Regressió lineal

Molts cops ens trobem al laboratori amb mesures de dues magnituds relacionades, de les quals hem de trobar la relació. La forma més simple de relació entre dues magnituds és la relació lineal:

$$Y = mX + b. (6)$$

Suposem que al laboratori prenem mesures simultànies de dues magnitud relacionades (p. ex. la intensitat i el voltatge en un circuit de corrent continu), i tenim N valors de les magnituds X i Y, si aquestes magnituds estan relacionades per una funció lineal (6), aleshores el pendent m i l'ordenada a l'origen b es poden trobar mitjançant les relacions:

$$m = \frac{\langle XY \rangle - \langle X \rangle \langle Y \rangle}{\langle X^2 \rangle - \langle X \rangle^2} = \frac{N \cdot \left(\sum_{i=1}^N x_i y_i\right) - \left(\sum_{i=1}^N x_i\right) \left(\sum_{i=1}^N y_i\right)}{\Delta} , \tag{7}$$

$$b = \frac{\langle X^2 \rangle \langle Y \rangle - \langle X \rangle \langle XY \rangle}{\langle X^2 \rangle - \langle X \rangle^2} = \frac{\left(\sum_{i=1}^N x_i^2\right) \cdot \left(\sum_{i=1}^N y_i\right) - \left(\sum_{i=1}^N x_i\right) \left(\sum_{i=1}^N x_i y_i\right)}{\Delta} . \tag{8}$$

on definim

$$\Delta \equiv \left| \begin{array}{cc} N & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \end{array} \right| = N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2$$

Evidentment és possible que les dades experimentals no s'ajustin a un comportament lineal; el paràmetre que ens mesura el grau d'acostament a una recta és el $factor\ de$ correlació

$$r^{2} = \frac{\left[\langle XY \rangle - \langle X \rangle \langle Y \rangle\right]^{2}}{\left[\langle X^{2} \rangle - \langle X \rangle^{2}\right] \cdot \left[\langle Y^{2} \rangle - \langle Y \rangle^{2}\right]} = \frac{\left[N \sum_{i=1}^{N} x_{i} y_{i} - \left(\sum_{i=1}^{N} x_{i}\right) \left(\sum_{i=1}^{N} y_{i}\right)\right]^{2}}{\left[N \sum_{i=1}^{N} x_{i}^{2} - \left(\sum_{i=1}^{N} x_{i}\right)^{2}\right] \cdot \left[N \sum_{i=1}^{N} y_{i}^{2} - \left(\sum_{i=1}^{N} y_{i}\right)^{2}\right]}$$

Es considera que un ajust és acceptable quan $r \ge 0.95$.

Exemple

Suposem que volem mesurar la constant d'elongació d'una molla, mesurant la força exercida sobre la molla i l'elongació que aquesta produeix. Sabem que la llei teòrica que relaciona aquestes variables és la llei de Hooke:

$$F = k(x - x_0) = kx - kx_0$$
,

on x és l'elongació i x_0 és la llargada en repòs. Suposem que fem 4 mesures de la força per 4 valors de x diferents:

$$F = \{0.5, 1.4, 1.9, 2.1\} \pm 0.2 \,\text{N}$$

$$x = \{10, 15, 20, 25\} \pm 1 \text{cm}$$

aleshores

$$\langle x \rangle = 17.5 \,\text{cm}, \qquad \langle x^2 \rangle = 337.5 \,\text{cm}^2,$$
$$\langle F \rangle = 1.475 \,\text{N}, \qquad \langle F^2 \rangle = 2.5575 \,\text{N}^2,$$
$$\langle xF \rangle = 29.125 \,\text{N} \,\text{cm},$$

el valor del pendent és

$$m = k = 0.106 \text{N/cm}$$
,

i l'ordenada a l'origen és

$$b = -k x_0 = -0.38$$
N.

El valor del factor de correlació

$$r = 0.958893$$

ens indica que les dades experimentals realment s'ajusten a la recta. A la figura 1 podem veure la representació gràfica dels punts experimentals i de la recta de regressió.

Ens faltaria per últim trobar l'error en els dos paràmetres m i b de la regressió. Aquest càlcul és força complicat, encara que sota certes condicions⁴ l'error de m i b es pot expressar en fórmules senzilles:

$$\epsilon_m^2 = \frac{\sigma^2}{\Delta} N$$

$$\epsilon_b^2 = \frac{\sigma^2 \langle X^2 \rangle}{\Delta} = \frac{\sigma}{\Delta} \sum_{i=1}^N X_i^2 ,$$

on hem utilitzat la incertesa mitjana del conjunt de mesures:

$$\sigma^2 = \frac{1}{N-2} \sum_{i=1}^{N} (Y_i - (mX_i + b))^2.$$

⁴Principalment, que els errors en la variable independent X siguin negligibles i que les mesures de la variable Y s'hagin fet amb el mateix instrument, de manera que l'error instrumental de les diferents mesures és el mateix.

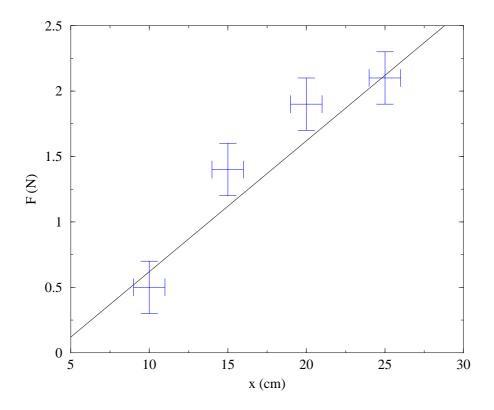


Figura 1: Representació de la força en funció de la distància d'una molla. Hi ha representats els punts experimentals (amb el seu error) així com la recta de regressió que més s'ajusta a les dades.

Apèndix A. Propagació d'Errors

Fórmula general

Recordeu que l'error en una magnitud derivada F que depèn de diverses magnituds x_j (j = 1, 2 ... n) segons una funció $F = f(x_1, ..., x_n)$, on cada magnitud x_j té un error diferent ϵ_j , ve donat per la següent expressió:

$$\epsilon_F = \sqrt{\sum_{j=1}^n \left(\frac{\partial F}{\partial x_j} \epsilon_j\right)^2} \,, \tag{9}$$

on $\partial F/\partial x_j$ és la derivada de la funció F respecte la variable x_j mantenint les altres variables constants.

Convé estudiar aquesta relació per a alguns casos simples.

Suma

Suposem que tenim una funció que depèn de dues variables, x i y, de la manera següent:

$$F = x + y \tag{10}$$

Segons la fórmula (9), l'error en F que arrosseguem per culpa de l'error en x i de l'error en y és:

$$\epsilon_F = \sqrt{1^2 \cdot \epsilon_x^2 + 1^2 \cdot \epsilon_y^2} = \boxed{\sqrt{\epsilon_x^2 + \epsilon_y^2}}.$$
(11)

Si enlloc d'una suma tenim una diferència el resultat és exactament el mateix.

Producte

Si F és una funció de x i de y del tipus $F = x \cdot y$, l'error en F vindrà donat, aplicant (9), per

$$\boxed{\epsilon_F = \sqrt{y^2 \epsilon_x^2 + x^2 \epsilon_y^2}}.$$
 (12)

Divisió

Si F depèn de x i y segons F = x/y, l'error en F serà:

$$\epsilon_F = \sqrt{\left(\frac{1}{y}\right)^2 \epsilon_x^2 + \left(-\frac{x}{y^2}\right)^2 \epsilon_y^2} \,. \quad (13)$$

Fórmules específiques de cada pràctica

Mesura de la Viscositat

• Error en el volum de l'esfera $V = \frac{4}{3}\pi r^3$:

$$\epsilon_V = \sqrt{(4\pi r^2)^2 \epsilon_r^2} = 4\pi r^2 \epsilon_r .$$

• Error en la densitat de l'esfera $\rho_{es} = m/V$:

$$\epsilon_{\rho_{es}} = \sqrt{\left(\frac{1}{V}\right)^2 \epsilon_m^2 + \left(-\frac{m}{V^2}\right)^2 \epsilon_V^2} .$$

• Error en la **velocitat mitjana**, $v_m = h/t$ (apartat 3 del full de respostes):

$$\epsilon_{v_m} = \sqrt{\left(\frac{1}{t}\right)^2 \epsilon_h^2 + \left(-\frac{h}{t^2}\right)^2 \epsilon_t^2} .$$

• Error en la **velocitat límit**, $v_{\text{lim}} = (1 + 2.4D/\Phi)v_m$ (apartat 4 del full de respostes). Considerem que Φ , el diàmetre del tub, té un error negligible.

$$\epsilon_{v_{\text{lim}}} = \sqrt{\left(1 + 2.4 \frac{D}{\Phi}\right)^2 \epsilon_{v_m}^2 + \left(2.4 \frac{v_m}{\Phi}\right)^2 \epsilon_D^2}$$
.

• Error en la **viscositat**,

$$\eta = \frac{D^2(\rho_{es} - \rho_l)g}{18v_{\lim}}.$$

Negligim l'error en la densitat del líquid, ρ_l , i en l'acceleració de la gravetat g.

$$\epsilon_{\eta} = \sqrt{\left(2D\frac{(\rho_{es} - \rho_{l})g}{18v_{\text{lim}}}\right)^{2}\epsilon_{D}^{2} + \left(\frac{D^{2}g}{18v_{\text{lim}}}\right)^{2}\epsilon_{\rho_{es}}^{2} + \left(-\frac{D^{2}(\rho_{es} - \rho_{l})g}{18v_{\text{lim}}^{2}}\right)^{2}\epsilon_{v_{\text{lim}}}^{2}}.$$

Mesura de densitats

• Error en el factor de correcció, $f = \rho_a/\rho_a'$. Negligim l'error en la densitat real de l'aigua:

$$\epsilon_f = \sqrt{\left(-\frac{\rho_a}{{\rho'_a}^2}\right)^2 \epsilon_{\rho'_a}^2} = \frac{\rho_a}{{\rho'_a}^2} \cdot \epsilon_{\rho'_a}.$$

• Error en la densitat de la llet, $\rho_{\text{llet}} = f \cdot \rho_{\text{mesurada}}$:

$$\epsilon_{\rho_{\rm llet}} = \sqrt{\rho_{\rm mesurada}^2 \cdot \epsilon_f^2 + f^2 \cdot \epsilon_{\rho_{\rm mesurada}}^2} \; .$$

La llei d'Ohm

• L'error en la **resistència**, R = V/I:

$$\varepsilon_R = \sqrt{\left(\frac{1}{I}\,\varepsilon_V\right)^2 + \left(-\frac{V}{I^2}\varepsilon_I\right)^2}$$

• Error en la resistència equivalent d'un circuit format per dues resistències en sèrie, $R_{\text{eq,ser}} = R_1 + R_2$. Es tracta d'un cas descrit per la fórmula (10):

$$\epsilon_{R_{\rm eq,ser}} = \sqrt{\epsilon_{R_1}^2 + \epsilon_{R_2}^2}$$
 .

• Error en la resistència equivalent d'un circuit format per dues resistències en paral·lel, $R_{\rm eq,par}^{-1}=R_1^{-1}+R_2^{-1}$.

$$\epsilon_{R_{\rm eq,par}} = \frac{1}{(R_1 + R_2)^2} \sqrt{(R_2^2 \epsilon_{R_1})^2 + (R_1^2 \epsilon_{R_2})^2} \; . \label{epsilon}$$