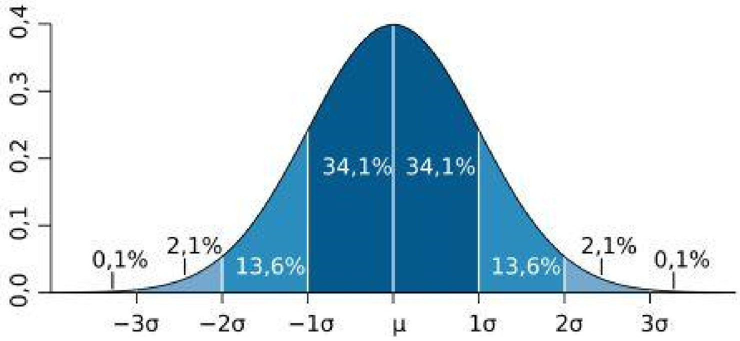
Iniciem els apunts amb el teorema del valor central i la propagació de variàncies. Fem desprès una deducció elemental de l'ajust lineal de mínim quadrats, reescrivim els resultats en forma matricial, cosa que fa immediata la generalització a regressió multilineal i no lineal. Uns apèndixs addicionals fa el text auto-contingut

Apunts breus sobre regressió lineal i no lineal

Josep Planelles

1.Teorema del límit central

El Teorema del límit central (o Teorema central del límit) diu que la distribució de la mitjana de qualsevol variable aleatori independent serà normal, si la mida de la mostra és prou gran.

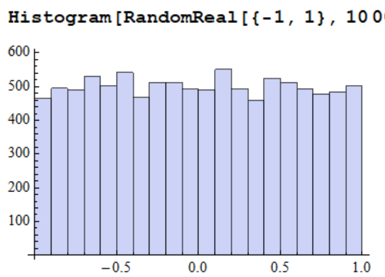


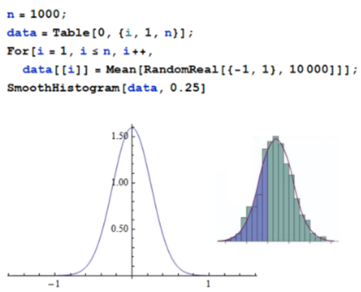
$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Distribució Gaussiana
o distribució normal

En altres paraules, si considerem una variable física que determinem experimentalment i imaginem que el seu valor "real" (lliure de tota imprecisió experimental) fos *a*, mentre que el seu valor experimental resultat de l'i-èssim experiment és $y_i = a + x_i$. Vol dir açò que la variable purament aleatòria és x_i i que el seu valor "real" hauria de ser zero. Doncs bé, el teorema del límit central ve a dir que, efectivament, el valor de la variable x és zero i el doble de la seua variància (que és la mateixa que la variància de la variable $y = a + x$, atès que *a* és un valor fix) ens dona un interval d'imprecisió que permet afirmar amb *certesa estadística* que *a* està dins de l'interval y ± 2σ, (per *certesa estadística* entenem donar un resultat amb una probabilitat del 95% de no equivocar-nos).

No demostrarem ací aquest teorema però presentem una simulació feta amb Mathematica:





La figura de l'esquerra presenta un histograma de la distribució de 10.000 nombres aleatoris entre els valors zero i la unitat. Com podem observar estan distribuïts més o menys uniformement en tot l'interval. A la dreta hem calculat la mitjana de 1000 distribucions de 10.000 nombres aleatoris distribuïts entre zero i un. Els resultats els hem emmagatzemat en la llista data, la qual és representada a la dreta amb un histograma suavitzat (junt amb un esquema que mostra un histograma i la seua suavització).

En resum, aquest teorema fonamenta l'ús de la mitjana $\bar{x}$ d'una mostra com la millor proposta del valor "real" μ de la variable i l'estimació de la seua precisió com aproximadament el doble de la variància (també anomenada distribució quadràtica): $\mu = \bar{x} \pm 2\sigma$.

Si volem ser *exactes*, i donar el resultat amb una probabilitat del 95%, és a dir, amb un àrea de valor 0.95 davall la campana de probabilitat, escriurem: $\mu = \bar{x} \pm 1.98\,\sigma$. Ara bé, aquesta campana implica el límit infinit d'una mostra gran. Com les mostres sempre són finites hom usa sovint el valor 2σ. Si la mostra és molt reduïda cal usar valors majors per assegurar la probabilitat de 95% (per exemple la distribució de Student assigna un valor 2.57 a mostres de 5 elements).

2.Propagació de la variància (o de la imprecisió)

Sovint mesurem una variable x, però ens interessa saber el valor de la magnitud $y = f(x)$. Aleshores sorgeix la pregunta: és la funció de la mitjana $f(\bar{x})$ la millor proposta per a y? Quina és la seua imprecisió?

A partir de $y = f(x)$ és immediat dir $dy = \left(\frac{df}{dx}\right) dx$. Aleshores, si identifiquem dx, dy amb les diferències finites $(x - x_i),\ (y - y_i)$ entre el valor "exacte" i la seua determinació experimental que inclou un petit error aleatori $\delta x, \delta y$, podem escriure:

$$(y - y_i)^2 = \left(\frac{df}{dx}\right)^2 (x - x_i)^2 \rightarrow \sum_i (y - y_i)^2 = \left(\frac{df}{dx}\right)^2 \sum_i (x - x_i)^2$$

que dividint pel nombre de mesures (N) en proporciona la relació entre les desviacions quadràtiques o variàncies:

$$S_y^2 = \left(\frac{df}{dx}\right)^2 S_x^2$$

2.1 Imprecisió de la mitjana

En el primer apartat hem donat arguments per triar la mitjana d'una mostra com la millor proposta per al valor de la variable aleatòria que es mesura. Ací en donarem un addicional: la seua imprecisió és N vegades menor que la de qualsevol element de la mostra dels N experiments. En efecte, escrivim $\bar{x} = \frac{1}{N}\sum_i x_i$. Per tant, la mitjana és una funció dels N valors aleatoris que constitueix la mostra: $\bar{x} = f(x_1, x_2, \dots, x_N)$. Podem calcular la seua derivada, $\frac{\partial \bar{x}}{\partial x_i} = \frac{1}{N}$, idèntica per a tots el valors x_i. Per tant, $S_{\bar{x}}^2 = \frac{1}{N^2}\sum_i S_x^2 = \frac{1}{N^2} N S_x^2 = \frac{S_x^2}{N}$.

$$\boxed{S_{\bar{x}}^2 = \frac{1}{N} S_x^2}$$

2.2 Propagació en cas d'una funció de dues variables

Considerem $z = f(x, y)$. Seguint el mateix raonament, $dz = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy$, és a dir,

$$\left(z - z_{ij}\right) = \left(\frac{\partial f}{\partial x}\right)_y (x - x_i) + \left(\frac{\partial f}{\partial y}\right)_x (y - y_j),$$

$$\rightarrow \frac{1}{N_i N_j} \sum_{ij} (z - z_{ij})^2 =$$

$$= \frac{1}{N_i N_j}\left(\frac{\partial f}{\partial x}\right)_y^2 N_j \sum_i (x - x_i)^2 + \frac{1}{N_i N_j}\left(\frac{\partial f}{\partial y}\right)_x^2 N_i \sum_i (y - y_i)^2 + \frac{2}{N_i N_j}\left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial f}{\partial y}\right)_x \sum_i (x - x_i) \sum_j (y - y_j)$$

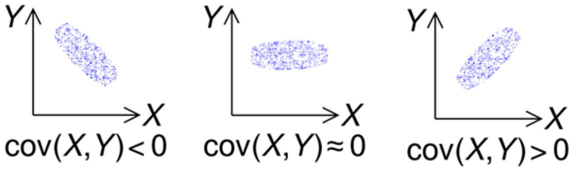
Que reescrivim: $S_z^2 = \left(\frac{\partial f}{\partial x}\right)_y^2 S_x^2 + \left(\frac{\partial f}{\partial y}\right)_x^2 S_y^2 + 2\left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial f}{\partial y}\right)_x S_{xy}^2 = \left[\left(\frac{\partial f}{\partial x}\right)_y \quad \left(\frac{\partial f}{\partial y}\right)_x\right] \begin{pmatrix} S_x^2 & S_{xy}^2 \\ S_{yx}^2 & S_x^2 \end{pmatrix} \begin{bmatrix} \left(\frac{\partial f}{\partial x}\right)_y \\ \left(\frac{\partial f}{\partial y}\right)_x \end{bmatrix}$,

on la matriu que inclou les variàncies s'anomena matriu de covariància.

Si les variables x i y són independents, la covariància S_{xy}^2 resulta ser zero, atès que els valors x_i són independents del valor y_i, per tant: $\sum_i(\bar{x}-x_i)\sum_j(\bar{y}-y_j) = N(x-\bar{x})N(y-\bar{y}) \approx 0$, perquè que la mitjana és (aproximadament) igual al valor "real" de la variable. Ara bé, si les variables aleatòries estan relacionades entre sí, el valor x_i condiciona el valor y_i i per tant ja no podem calcular les sumes com independents, atès que no ho són. Si escrivim:

$$\frac{1}{N^2}\sum_{ij}(\bar{x}-x_i)(\bar{y}-y_j) = \frac{1}{N^2}\sum_{ij}x_iy_j - \frac{1}{N}\,\bar{y}\sum_i x_i - \frac{1}{N}\,\bar{x}\sum_j y_i + \bar{x}\,\bar{y} = (\overline{xy}-\bar{x}\,\bar{y})$$

En altres paraules, la mitjana dels productes no és igual al producte de les mitjanes. En l'apartat següent demostrarem que $(\overline{xy}-\bar{x}\,\bar{y})$ és proporcional al pendent del ajust lineal de $y = f(x)$.



A l'apèndix 7.1 donem més detalls sobre aquest paràmetre estadístic.

3.Ajust Lineal de Mínim Quadrats

Anomenem y_i^c als valors de la variable aleatòria que vénen donats per la regressió, $y_i^c = A + Bx_i$. Volem trobar els valors A y B que fan que la recta passe pel mig del núvol de punts experimentals. En altres paraules, volem trobar els valors A i B que facin mínima la distància dels punts experimentals a la recta proposada, és a dir, volem que la magnitud D_{reg}^2, que definim com a suma de les distàncies dels punts a la recta,

$$D_{reg}^2 = \sum_{ij}(y_{ij} - y_i^c)^2 = \sum_{ij}(y_{ij} - A - Bx_i)^2,$$

sigui mínima. Fem la suma de diferències elevades al quadrat, en lloc de la suma de diferències, a fi d'evitar que diferències negatives resten en lloc de sumar.

Si triem *A* i *B* de manera que la suma de quadrats siga mínima, aleshores, les derivades respecte d'A i B seran nul·les,

$$\frac{\partial D_{reg}^2}{\partial A} = 2\sum_{ij}(y_{ij} - A - Bx_i) = 0 \quad \Rightarrow N\,\bar{y} - N\,A - N\,B\,\bar{x} = 0$$

$$\frac{\partial D_{reg}^2}{\partial B} = 2\sum_{ij}(y_{ij} - A - Bx_i)x_i = 0 \Rightarrow N\,\overline{xy} - N\,A\,\bar{x} - N\,B\,\overline{x^2} = 0$$

on: $\quad \bar{y} = \frac{1}{N}\sum_{ij} y_{ij}, \quad \overline{xy} = \frac{1}{N}\sum_{ij} x_i y_{ij}, \quad$ etc.

Des d'aquestes equacions obtenim els valors A i B que minimitzen la distància dels punts experimentals a la recta proposada i que és immediat trobar. Els resultats són:

$$B = \frac{\overline{\mathrm{xy}} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2}\ ; \qquad A = \bar{y} - B\ \bar{x}$$

Adonem-nos que el numerador del pendent B és precisament la covariància S_{xy}^2.

La segona tasca a abordar és la determinació de la bondat de l'ajust i la dels paràmetres A i B. Una mesura d'aquesta bondat ens la dóna la suma de quadrats, si la normalitzem dividint-la pel nombre de punts. Aquesta normalització la fem per evitar que, en afegir punts experimentals, creixca el valor de la suma de quadrats, cosa que faria la impressió d'empitjorar la bondat de l'ajust en fer més mesures, cosa completament absurda. De la mateixa manera, en comptes de dividir-la pel nombre de punts N, dividirem per N-2. Açò ho fem perquè un experiment en què només hem fet dues mesures no ens doni un error zero (per dos punts sempre passa una recta). En realitat, si només hem fet dues mesures no tenim ni idea de quin és l'error del nostre experiment, i.e. la bondat de la configuració és completament indeterminada. Podem fer que les matemàtiques ens diguin que la bondat és indeterminada si definim el paràmetre de bondat com la següent *dispersió*:

$$S_{reg}^2 = \frac{D_{reg}^2}{N-2} = \frac{1}{N-2}\sum_{ij}(y_{ij} - y_i^c)^2 = \frac{1}{N-2}\sum_{ij}(y_{ij} - A - Bx_i)^2$$

En efecte, si només hem fet dues mesures, la recta ajustada passa pels únics dos punts experimentalment determinats, per tant, la distància dels punts experimentals a la recta és zero i també ho és la suma de quadrats. Ara bé, el denominador (N-2) del paràmetre de bondat també és zero ja que N=2 i, en matemàtiques, una fracció zero dividit per zero està indeterminada.

Tenim doncs un primer paràmetre que ens indica la bondat de l'ajust, però cal ser més concrets i determinar la imprecisió dels paràmetres ajustats. Comencem per B, que podem reescriure com:

$$B = \frac{1}{\overline{x^2} - \bar{x}^2}\frac{1}{N}\sum_{ij}(x_i - \bar{x})y_{ij}$$

En aquesta expressió només y_{ij} és aleatori i per tant presenta dispersió i que anomenem *dispersió de la regressió*. Com que B depèn de y_{ij} i y_{ij} té un error aleatori, també B presentarà un error aleatori. La propagació de la dispersió dóna lloc a:

$$S_B^2 = \frac{1}{\left[\overline{x^2} - \bar{x}^2\right]^2}\frac{1}{N^2}\sum_{ij}(x_i - \bar{x})^2 S_{reg}^2 = \frac{S_{reg}^2}{\left[\overline{x^2} - \bar{x}^2\right]^2}\frac{1}{N^2}N(\overline{x^2} - \bar{x}^2) = \frac{S_{reg}^2}{N(\overline{x^2} - \bar{x}^2)} = \frac{S_{reg}^2}{D_x^2},$$

on hem introduït la notació: $D_x^2 = \sum_i (x_i - \bar{x})^2 = N(\overline{x^2} - \bar{x}^2)$.

De manera anàloga, el paràmetre $A = \bar{y} - B\bar{x} = \sum_{ij}\left(\frac{1}{N} - \frac{\bar{x}\,(x_i - \bar{x})}{D_x^2}\right)\, y_{ij}$, i com la suma $\sum_i (x_i - \overline{x})$ val zero, resulta una dispersió:

$$S_A^2 = \left(\frac{1}{N} + \frac{\bar{x}^2}{D_x^2}\right) S_{reg}^2 = \frac{\overline{x^2}}{D_x^2}\ \ S_{reg}^2,$$

on la darrera igualtat s'obté substituint el valor $D_x^2 = N(\overline{x^2} - \bar{x}^2)$, i ordenant l'expressió obtinguda.

Finalment, ens pot interessar calcular la dispersió d'un valor ajustat $y_i^c = A + Bx_i$. Podríem aplicar la propagació de la dispersió i escriure $S^2_{y_i^c} = S_A^2 + x_i^2 S_B^2$, però atenció açò no seria correcte perquè pendent i ordenada no són estadísticament independents.[1] Cal reescriure $y_i^c = \bar{y} - B(x_i - \bar{x})$ i ara sí, amb $S_{\bar{y}}^2 = \frac{S_y^2}{N} \equiv \frac{S_{reg}^2}{N}$, trobem:

$$S^2_{y_i^c} = \frac{S_{reg}^2}{N}\left[1 + \frac{N(x_i - \bar{x})^2}{D_x^2}\right]$$

El mateix resultat s'obté a partir de $y_i^c = A + Bx_i$ fent ús de la covariància S_{AB}^2 (vegeu apèndix, secció 7.1).

La imprecisió o error es pot estimar com aproximadament el doble de l'arrel quadrada de la dispersió (recordeu que calculàvem quadrats per evitar números negatius. Ara hem d'eliminar els quadrats). De manera més exacta, la imprecisió o error s'ha de calcular com a *t* vegades l'arrel quadrada de la dispersió, on el paràmetre *t*, anomenat t de Student, té en compte el nombre de graus de llibertat del nostre experiment (que és igual en aquest cas al nombre de mesures menys 2, com havíem comentat anteriorment).

4.Ajust Lineal de Mínim Quadrats en forma matricial

Podem escriure l'equació $y_i^c = a + bx_i$ de la recta[2] en forma matricial $[y_i^c] = [1\ x_i]\begin{bmatrix} a \\ b \end{bmatrix}$. Aquesta equació és certa per a una *i* qualsevol. Per tant, és valida per a totes les *is*, de manera que podem escriure $Y = X\,A$, on *Y* és un vector columna que conté totes les y_i, $X = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}$ i $A = \begin{bmatrix} a \\ b \end{bmatrix}$.[3]

Podem aïllar els paràmetres *A* multiplicant a esquerres per la trasposta d'*X*, de manera que el producte $X^T X$ és una matriu quadrada la qual pot ser invertida $(X^T X)^{-1}$, cosa que permet aïllar els paràmetres: $A = (X^T X)^{-1} X^T Y$.

4.1 Càlcul de les imprecisions: La covariància

En la pàgina 2 hem obtingut la propagació de la variància d'una funció de dues variables aleatòries $z = f(x, y)$:

$$S_z^2 = \left(\frac{\partial f}{\partial x}\right)_y^2 S_x^2 + \left(\frac{\partial f}{\partial y}\right)_x^2 S_y^2 + 2\left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial f}{\partial y}\right)_x S_{xy}^2 = \left[\left(\frac{\partial f}{\partial x}\right)_y \quad \left(\frac{\partial f}{\partial y}\right)_x\right]\begin{pmatrix} S_x^2 & S_{xy}^2 \\ S_{yx}^2 & S_y^2 \end{pmatrix}\begin{bmatrix} \left(\frac{\partial f}{\partial x}\right)_y \\ \left(\frac{\partial f}{\partial y}\right)_x \end{bmatrix} = Z^T \cdot cov \cdot Z,$$

on Z és el vector de derivades i cov la matriu de covariància. Considerem per exemple el cas d'ajust general $y_i^c = f(x_i, a, b)$ d'una funció de la variable de control x que depèn de dos paràmetres (a, b). Assumim que la

[1] La propagació de la dispersió de variables no independents inclou un terme funció la seua covariància (la covariància de variables estadísticament independents és zero i aquest terme desapareix).
[2] Hem escrit pendent i ordenada en minúscules perquè ara reservem les lletres capitals per a les matrius.
[3] De manera completament anàloga podem ajustar polinomis de major grau. A l'apèndix ho exemplifiquem al cas d'ajust a una paràbola.

variable aleatòria és y_i i la no-aleatòria o variable de control la x_i, de manera que en l'expressió $y_i^c = f(x_i, a, b)$ l'aleatorietat o imprecisió de la y_i se transmet a (o des de) el paràmetres de l'ajust a, b: $y_i(a, b)$.

Per tant podem escriure $S_{y_i}^2 = Z_i^T \cdot cov \cdot Z_i$, amb $Z_i^T = \left[\left(\frac{\partial y_i}{\partial a}\right)_b \quad \left(\frac{\partial y_i}{\partial b}\right)_a\right] \equiv [za_i \quad zb_i]$ i cov la matriu de covariància: $cov = \begin{pmatrix} S_a^2 & S_{ab}^2 \\ S_{ba}^2 & S_b^2 \end{pmatrix}$. Multiplicant a la dreta per Z_i^T obtenim: $S_{y_i}^2\, Z_i^T = Z_i^T \cdot cov \cdot Z_i \cdot Z_i^T$, que podem reescriure en la forma $Z_i^T (S_{y_i}^2 - cov \cdot Z_i \cdot Z_i^T) = 0$. És a dir, $S_{y_i}^2 = cov \cdot Z_i \cdot Z_i^T$. De manera explícita:

$$S_{y_i}^2 = cov \cdot \left(Z_i \cdot Z_i^T\right) = cov \cdot \begin{bmatrix} za_i \\ zb_i \end{bmatrix} \cdot [za_i\ zb_i] = cov \cdot \begin{pmatrix} za_i^2 & za_i\, zb_i \\ za_i\, zb_i & zb_i^2 \end{pmatrix}$$

En efectuar la suma de les i a l'esquerra de la igualtat trobem la dispersió de la regressió:[4] $S_r^2 = \sum_i S_{y_i}^2$. La suma de la dreta de la igualtat dóna lloc a:

$$cov \cdot \sum_i \begin{pmatrix} za_i^2 & za_i\, zb_i \\ za_i\, zb_i & zb_i^2 \end{pmatrix} = cov \cdot \left[\sum_i \begin{pmatrix} za_i^2 & za_i\, zb_i \\ za_i\, zb_i & zb_i^2 \end{pmatrix}\right] = cov \cdot \begin{pmatrix} \Sigma za_i^2 & \Sigma za_i\, zb_i \\ \Sigma za_i\, zb_i & \Sigma zb_i^2 \end{pmatrix} = cov \cdot (Z^T Z).$$ [5]

Per tant, trobem que la matriu de covariància se pot expressar en termes de la matriu de les derivades de la funció $f(x, a, b)$ respecte dels paràmetres i la S_r^2: $\quad cov = (Z^T Z)^{-1} S_r^2$.

Cal fer notar que encara que la demostració l'hem feta per al cas d'una variable de control i dos paràmetres, només rellegint-la veiem que seria idèntica si hi hagueren més variables de control. També seria la mateixa si hi ha més paràmetres, tan sols canvia la matriu Z, que per a dos paràmetres és una matriu $n \times 2$ per a tres és una matriu $n \times 3$ etc.

4.2 Càlcul de les imprecisions en ajustos lineals: La covariància en termes de la matriu X de la variable no aleatòria

En el cas lineal, $y = a + bx$, els paràmetres de l'ajust són a i b. Per tant, $Z_i^T = \left[\left(\frac{\partial y_i}{\partial a}\right)_b \quad \left(\frac{\partial y_i}{\partial b}\right)_a\right] = [1 \quad x_i]$, de manera que Z no és més que la matriu $X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$. Per tant, covariància resulta: $cov = (X^T X)^{-1} S_r^2$.

4.3 Algunes comprovacions

La lectura d'aquest apartat és innecessària. És just una altra forma de veure que $cov = (X^T X)^{-1} S_r^2$, a partir de l'anàlisi del contingut de la matriu $X^T X$:

$$X^T X = \begin{bmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$$

[4] De manera anàloga a $S_{\bar{x}}^2 = n S_{\bar{x}}^2$ tenim que $S_r^2 = n S_{\bar{y}}^2 = n \frac{1}{n} \sum_i S_{y_i}^2 = \sum_i S_{y_i}^2$.

[5] Adonem-nos que $\begin{pmatrix} \Sigma za_i^2 & \Sigma za_i\, zb_i \\ \Sigma za_i\, zb_i & \Sigma zb_i^2 \end{pmatrix}$ és la matriu (2×2) producte de $Z^T(2 \times n)$ per $Z(n \times 2)$.

Si calculem el determinant d'aquesta matriu trobem: $\det(X^T X) = n^2\left(\overline{x^2} - \bar{x}^2\right) = n\sum_i (x_i - \bar{x})^2 = nD_x^2$.

Atès que la inversa de la matriu $M=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ és: $M^{-1} = \frac{1}{\det(M)}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$ tenim que:

$$(X^T X)^{-1} = \frac{1}{D_x^2}\begin{bmatrix} \overline{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix}$$

Si mirem els termes diagonals observem que, excepte que falta multiplicar per S_r^2, són precisament les variàncies de pendent i ordenada (veure pàgina 4. També en apèndix 7.1). Els extra-diagonals són iguals entre si i, excepte el factor S_r^2, són la covariància S_{ab}^2 (calculada en l'apèndix 7.1). La matriu de covariància de la regressió és doncs $cov = (X^T X)^{-1} S_r^2$. Finalment, atès que $y_i^c = a + bx_i$, la variància de la y_i^c resulta:

$$S_{y_i^c}^2 = (1 \quad x_i)\ cov \begin{pmatrix} 1 \\ x_i \end{pmatrix} = (1 \quad x_i)(X^T X)^{-1}\begin{pmatrix} 1 \\ x_i \end{pmatrix} S_r^2.$$

5.Regressió Multi-Lineal que passa per l'origen

Si tenim dues variables, aleshores $[z_i^c] = [x_i\ y_i]\begin{bmatrix} a \\ b \end{bmatrix}$. En general, $Y = X\,A$, amb la mateixa definició de Y i A, amb $X = \begin{bmatrix} x_1 & y_1 \\ \cdots & \cdots \\ x_n & y_n \end{bmatrix}$. Com adés, podem aïllar els paràmetres $A = (X^T X)^{-1} X^T Y$ i determinar la matriu de covariància $(X^T X)^{-1}$, la diagonal de la qual representa les variàncies del paràmetres, mentre que el producte $(x_i\ y_i)(X^T X)^{-1}\begin{pmatrix} x_i \\ y_i \end{pmatrix} S_r^2$ són les $S_{y_i^c}^2$.

5.1 Regressió Multi-Lineal forçada a passar per l'origen

Considerem el problema amb dos variables: $\tilde{z}_i^c = \alpha + a\,x_i + b\,y_i$, on ($\alpha$, a, b) són paràmetres que determinem amb l'ajust multilineal. Escrivim:

$$[\tilde{z}_i^c] = [1 \quad x_i \quad y_i]\begin{bmatrix} \alpha \\ a \\ b \end{bmatrix} \quad \Leftrightarrow \quad \tilde{Z} = X \cdot A \quad \text{amb } \tilde{Z} = \begin{bmatrix} \tilde{z}_1^c \\ \cdots \\ \tilde{z}_N^c \end{bmatrix},\ X = \begin{bmatrix} 1 & x_1 & y_1 \\ \cdots & \cdots & \cdots \\ 1 & x_N & y_N \end{bmatrix},\ A = \begin{bmatrix} \alpha \\ a \\ b \end{bmatrix}.$$

Aleshores, $A = (X^T X)^{-1} X^T \tilde{Z}$, i la matriu cov de covariàncies: $cov = (X^T X)^{-1} S_r^2$.

Podríem, però, definir $z_i^c = \tilde{z}_i^c - \alpha$ i assumir que l'α òptim que trobem en l'ajust lineal és exactament, sense cap error, el valor que hauria de tindre aquest paràmetre (cosa que segurament no és certa i que equival a acumular la seua imprecisió en la dels altres paràmetres[6]). Aleshores podem escriure:

$$\tilde{z}_i^c - \alpha = z_i^c = [x_i \quad y_i]\begin{bmatrix} a \\ b \end{bmatrix} \quad \Leftrightarrow \quad \tilde{Z} = X \cdot A \quad \text{amb } \tilde{Z} = \begin{bmatrix} z_1^c \\ \cdots \\ z_N^c \end{bmatrix},\ X = \begin{bmatrix} x_1 & y_1 \\ \cdots & \cdots \\ x_N & y_N \end{bmatrix},\ A = \begin{bmatrix} a \\ b \end{bmatrix}.$$

[6] Se fa una cosa semblant quan en els ajustos *y=f(x)* assumim que *x* és la variable no aleatòria de control mentre que *y* és la variable aleatòria. Per exemple, en una cinètica de reacció, on mesurem concentració de reactiu en front del temps, considerem usualment que l'error del temps és rebutjable i, en fer açò, estem acumulant en l'error de la concentració el seu propi error més l'error en la mesura del temps.

Aleshores, $A = (X^T X)^{-1} X^T Z$. La matriu de covariàncies $cov = (X^T X)^{-1} S_r^2$ i variància de la funció calculada z_i^c resulta: $S_{z_c}^2 = [x_i \quad y_i]\,(X^T X)^{-1} \begin{bmatrix} x_i \\ y_i \end{bmatrix} S_r^2$. Finalment, com hem assumit que $\tilde{z}_i^c - \alpha = z_i^c$ amb α el valor *exacte*, trobem que $S_{z_c}^2 = S_{\tilde{z}_c}^2$.

6.Regressió no Lineal

Estudiem la relació no lineal $y = f(\underline{x}, \underline{\theta})$, amb variables $\underline{x} = (x_1, \ldots, x_n)$ i paràmetres $\underline{\theta} = (a, b, \ldots \theta_p)$. En aquest cas la relació entre y i $\underline{x}$ no és lineal de manera que no pot ser expressada en termes d'un sistema lineal d'equacions amb solució única. Per trobar el valor òptim del paràmetres cal fer ús de mètodes iteratius a partir d'un *initial guess* o punt inicial. Per exemple podem emprar el mètode de Newton o altres mètodes semblants. Excel conté la macro SOLVER que usa el mètode GRG (*Generalized Reduced Gradient*) per a trobar mínims. Aquest mètode, com el de Newton, fa ús de les derivades per a accelerar la convergència. Mathematica té el comandament FindRoot, que també usa mètodes tipus Newton.

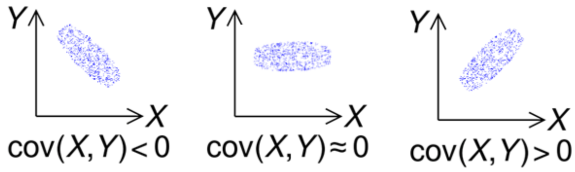
Per tant, el primer pas de l'ajust no lineal passa per minimitzar D_r^2 amb un mètode iteratiu. Una volta trobats el paràmetres òptims, calcularem la matriu Z de derivades $za_i = \left(\frac{\partial y_i}{\partial a}\right)_0$, $zb_i = \left(\frac{\partial y_i}{\partial b}\right)_0$... de la funció respecte dels paràmetres, per a cada valor $\underline{x_i}$ de les variables de control i el valor òptim $a, b,$ ...dels paràmetres.

Aleshores, calcularem la matriu de covariància $cov = (Z_0^T Z_0)^{-1} S_r^2$, les diagonals de la qual són les variàncies dels paràmetres i, amb d'ella calcularem la variància de la y calculada, $S_{y_i^c}^2 = Z_i^T \, cov \, Z_i$, on $Z_i = \begin{pmatrix} Z_{i1}^0 & \ldots & Z_{ip}^0 \end{pmatrix}$ és la fila de derivades calculades, com hem dit abans, per al valor òptim dels paràmetres.

7.Apèndix

7.1 Covariància

La covariància de dues variables aleatòries *x* i *y* és: $cov(x,y) \equiv S_{xy}^2 = \frac{1}{n}\Sigma(x_i-\bar{x})(y_i-\bar{y})$.



Una covariància positiva/negativa significa una relació directa/inversa de dos variables. La covariància d'una variable amb ella mateixa se diu variància $cov(x,x) \equiv S_x^2 = \frac{1}{n}\Sigma(x_i-\bar{x})^2$. De la definició, la covariància de constants per variables resulta: $cov(ax,by) = a \cdot b\; cov(x,y)$.

Si tenim dues magnitud *a* i *b* que són funció lineal d'una mateixa variable aleatòria *y*, $a = \Sigma a_i y_i$, $b = \Sigma b_i y_i$, aleshores $cov(a,b) = \Sigma a_i b_i\, S_y^2$.

En els apunts hem fet ús de la independència estadística de la mitjana $\bar{y}$ i la pendent, que ara comprovem després de reescriure: $\bar{y} = \sum_i y_i \frac{1}{n}$; $\quad b = \frac{1}{x^2-\bar{x}^2}\frac{1}{N}\sum_i(x_i-\bar{x})y_i = \sum_i y_i \frac{(x_i-\bar{x})}{D_x^2}$.

Aleshores, $cov(\bar{y},b) = \sum_i \frac{1}{n}\frac{(x_i-\bar{x})}{D_x^2}\quad S_y^2 = \frac{S_y^2}{n\,D_x^2}\sum_i(x_i-\bar{x}) = 0$.

Finalment, calcularem la covariància de la pendent i l'ordenada:

Escrivim: $\quad b = \sum_i y_i \frac{(x_i-\bar{x})}{D_x^2} \qquad\qquad a = \bar{y} - b\;\bar{x} = \sum_i y_i \frac{1}{n} - \bar{x}\sum_i y_i \frac{x_i-\bar{x}}{D_x^2} = \sum_i y_i\, [\frac{1}{n} - \bar{x}\frac{x_i-\bar{x}}{D_x^2}]$

Aleshores, si anomenem $A_i = \frac{x_i-\bar{x}}{D_x^2}$, tenim que:

$$cov(a,b) = S_y^2 \sum_i A_i \left[\frac{1}{n} - \bar{x}\, A_i\right] = S_y^2\, (\frac{1}{n}\sum_i A_i - \bar{x}\sum_i A_i^2) = -\,S_y^2\,\bar{x}\sum_i\left(\frac{x_i-\bar{x}}{D_x^2}\right)^2$$

$$\boxed{cov(a,b) = -\frac{\bar{x}}{D_x^2}\, S_y^2}$$

Exercici: calculeu $S_{y_i^c}^2$ a partir de l'expressió $y_i^c = A + Bx_i$ tenint en compte que *A* i *B* no són estadísticament independents. Dades: $S_A^2 = \frac{\overline{x^2}}{D_x^2} S_r^2$, $S_B^2 = \frac{S_r^2}{D_x^2}$, $S_{AB}^2 = -\frac{\bar{x}}{D_x^2} S_r^2$.

Des de l'apèndix 2, $S_{y_i^c}^2 = (\frac{\partial y_i^c}{\partial A})_B^2\, S_A^2 + (\frac{\partial y_i^c}{\partial B})_A^2\, S_B^2 + 2(\frac{\partial y_i^c}{\partial A})_B(\frac{\partial y_i^c}{\partial B})_A S_{AB}^2$.

Substituint: $\quad S_{y_i^c}^2 = \frac{\overline{x^2}}{D_x^2} S_r^2 + x_i^2\,\frac{S_r^2}{D_x^2} - 2x_i \frac{\bar{x}}{D_x^2} S_r^2 = \frac{S_r^2}{D_x^2}\left(\overline{x^2} + x_i^2 - 2\bar{x}x_i\right)$

$$= \frac{S_r^2}{D_x^2}\left(\overline{x^2} + \bar{x}^2 - \bar{x}^2 + x_i^2 - 2\bar{x}x_i\right) = \frac{S_r^2}{D_x^2}\left[(x_i-\bar{x})^2 + \frac{D_x^2}{n}\right]$$

$$\rightarrow \quad \boxed{S_{y_i^c}^2 = \frac{S_r^2}{n}[1 + \frac{n(x_i-\bar{x})^2}{D_x^2}]}$$

7.2 Ajust a una paràbola

Podem escriure l'equació $y_i^c = a + bx_i + cx_i^2$ de la paràbola en forma matricial $[y_i^c] = \begin{bmatrix} 1 \; x_i \; x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Aquesta equació és certa per a qualsevol i. Per tant, és valida per a totes les is, de manera que podem escriure $Y = X\,A$, amb $Y = \begin{bmatrix} y_1 \\ \cdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \dots & \dots & \dots \\ 1 & x_n & x_n^2 \end{bmatrix}$ i $A = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Podem calcular els paràmetres A multiplicant a esquerres per la trasposta d'X, de manera que el producte $X^T X$ és una matriu quadrada que pot ser invertida $(X^T X)^{-1}$, cosa que permet aïllar els paràmetres: $A = (X^T X)^{-1} X^T Y$.

A partir de la matriu $X^T X$:

$$X^T X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ x_1^2 & \cdots & x_n^2 \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_1^2 \\ \dots & \dots & \dots \\ 1 & x_n & x_n^2 \end{bmatrix} = \begin{bmatrix} n & \Sigma x_i & \Sigma x_i^2 \\ \Sigma x_i & \Sigma x_i^2 & \Sigma x_i^3 \\ \Sigma x_i^2 & \Sigma x_i^3 & \Sigma x_i^4 \end{bmatrix},$$

i de manera anàloga (però més feixuga) al cas lineal es pot comprovar que la matriu de covariància de la regressió és també: $cov = (X^T X)^{-1} S_r^2$. De fet, aquesta expressió és vàlida per tots als ajustos polinòmics.

Finalment, atès que les derivades de y_i^c respecte a, b, c són, respectivament, $1, x_i, x_i^2$, la variància $S_{y_i^c}^2$ resulta:

$$S_{y_i^c}^2 = (1\; x_i\; x_i^2)\; cov \begin{pmatrix} 1 \\ x_i \\ x_i^2 \end{pmatrix} = (1\; x_i\; x_i^2)(X^T X)^{-1} \begin{pmatrix} 1 \\ x_i \\ x_i^2 \end{pmatrix} S_r^2.$$

8.Excercicis

Ajust lineal

Versió amb excel

[1	x]	y
1	1	2,2
1	2	3,3
1	3	3,8
1	4	5,2
1	5	6,2
1	6	6,8
1	7	8,1
1	8	9,3
1	9	9,9
1	10	10,6

$x^T X$	10	55
	55	385

$x^T Y$	65,4
	439

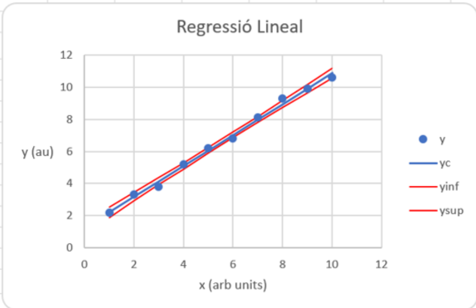
$(x^T X)^{-1} x^T Y=$	a	1,253333
	b	0,961212

x]	y	y_c	$(y-y_c)^2$
1	2,2	2,214545	0,00021157
2	3,3	3,175758	0,01543618
3	3,8	4,13697	0,11354858
4	5,2	5,098182	0,01036694
5	6,2	6,059394	0,01977006
6	6,8	7,020606	0,04866703
7	8,1	7,981818	0,01396694
8	9,3	8,94303	0,12742736
9	9,9	9,904242	1,7998E-05
10	10,6	10,86545	0,07046612
		Sr^2	0,05248485

cov	0,0244929	-0,0035
	-0,003499	0,000636

	estimació	st error	{interval de confiança}	
a	1,253333	0,156502	0,89181332	1,6148533
b	0,961212	0,025223	0,9029479	1,0194763

S_{yc}	y_{inf}	y_{sup}
0,1346519	1,9035	2,525591
0,1142002	2,911955	3,43956
0,0960448	3,915106	4,358833
0,0817306	4,909384	5,286979
0,0735359	5,889526	6,229262
0,0735359	6,850738	7,190474
0,0817306	7,793021	8,170616
0,0960448	8,721167	9,164894
0,1142002	9,64044	10,16804
0,1346519	10,55441	11,1765



Versió amb Mathematica (mateix problema)

```
ClearAll["Global`*"];
dadesxy = {{1, 2.2}, {2, 3.3}, {3, 3.8}, {4, 5.2}, {5, 6.2}, {6, 6.8}, {7, 8.1}, {8, 9.3}, {9, 9.9}, {10, 10.6}};
n = Length[dadesxy];
tStudent = 2.31;

Print["Solució paràmetres amb resolució sistema equacions"];
D2[a0_, a1_] = Sum[(dadesxy[[i, 2]] - a0 - a1 * dadesxy[[i, 1]])^2, {i, 1, n}];
eq1 = D[D2[a0, a1], a0]; eq2 = D[D2[a0, a1], a1];
sol = Solve[{eq1 == 0, eq2 == 0}, {a0, a1}];
Print["b0 = ", b0 = sol[[1, 1, 2]], "   b1 = ", b1 = sol[[1, 2, 2]]];
Print["sr2 = ", sr2 = D2[b0, b1]/(n - 2)];
Print[" "];
Print["Solució matricial"];
```

$$(* \; Y = X\,A \quad ; \quad X = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \quad ; \quad A = \begin{bmatrix} a \\ b \end{bmatrix} \;\Longrightarrow\; A = (X^T X)^{-1} X^T Y \; *)$$

```
X = {}; Y = {};
For[i = 1, i ≤ n, i++,
  X = AppendTo[X, {1, dadesxy[[i, 1]]}];
  Y = AppendTo[Y, dadesxy[[i, 2]]];
 ];
```

```
Print["{b0,b1} = ", Inverse[Transpose[X].X].Transpose[X].Y];
Print[" "];
Print["Variàncies-covariàncies dels paràmetres"];
cov = Inverse[Transpose[X].X] sr2;
Print[cov // MatrixForm];
Print["s(b0) = ", √cov[[1, 1]], ";    s(b1) = ", √cov[[2, 2]]];
Print[" "];
Print["Imprecissions de la y_calculada"];
```

$$(*\quad S^2_{y_i^c} = (1 \quad x_i)(X^T X)^{-1}\begin{pmatrix}1\\ x_i\end{pmatrix} S_r^2 \quad *)$$

```
s2yc = {};
For[i = 1, i ≤ n, i++,
  s2yc = AppendTo[s2yc, {{1, X[[i, 2]]}}.cov.X[[i]]];
 ];
taula = {{"y", "ycalc", "Syc", " ", "interval", "de", " confiança", " "}};
For[i = 1, i ≤ n, i++,
 taula = AppendTo[taula, {dadesxy[[i, 2]], aux = b0 + b1 * dadesxy[[i, 1]], aux2 = √s2yc[[i, 1]], "{",
     aux - aux2 * tStudent, ",", aux + aux2 * tStudent, "}"}];
]; taula // TableForm
```

En fer intro, el output que resulta és:

```
Solució paràmetres amb resolució sistema equacions
b0 = 1.25333    b1 = 0.961212
sr2 = 0.0524848

Solució matricial
{b0,b1} = {1.25333, 0.961212}

Variàncies-covariàncies dels paràmetres
( 0.0244929    -0.00349899 )
( -0.00349899   0.00063618 )
s(b0) = 0.156502;    s(b1) = 0.0252226

Imprecissions de la y_calculada
```

y	ycalc	Syc		interval	de	confiança	
2.2	2.21455	0.134652	{	1.9035	,	2.52559	}
3.3	3.17576	0.1142	{	2.91196	,	3.43956	}
3.8	4.13697	0.0960448	{	3.91511	,	4.35883	}
5.2	5.09818	0.0817306	{	4.90938	,	5.28698	}
6.2	6.05939	0.0735359	{	5.88953	,	6.22926	}
6.8	7.02061	0.0735359	{	6.85074	,	7.19047	}
8.1	7.98182	0.0817306	{	7.79302	,	8.17062	}
9.3	8.94303	0.0960448	{	8.72117	,	9.16489	}
9.9	9.90424	0.1142	{	9.64044	,	10.168	}
10.6	10.8655	0.134652	{	10.5544	,	11.1765	}

```
(* Fent ús de LinearModelFit *)
sol = LinearModelFit[dadesxy, {1, x}, x];
Print["Covariance =", sol["CovarianceMatrix"] // MatrixForm, ";  ", "sr2 =", sol["EstimatedVariance"],
 ";  ", sol["ParameterConfidenceIntervalTable"], ";    ", sol["MeanPredictionConfidenceIntervalTable"]]
```

```
Covariance =( 0.0244929    -0.00349899 ); sr2 =0.0524848;
            ( -0.00349899   0.00063618 )
```

	Estimate	Standard Error	Confidence Interval
1	1.25333	0.156502	{0.892439, 1.61423};
x	0.961212	0.0252226	{0.903049, 1.01938}

Observed	Predicted	Standard Error	Confidence Interval
2.2	2.21455	0.134652	{1.90404, 2.52505}
3.3	3.17576	0.1142	{2.91241, 3.4391}
3.8	4.13697	0.0960448	{3.91549, 4.35845}
5.2	5.09818	0.0817306	{4.90971, 5.28665}
6.2	6.05939	0.0735359	{5.88982, 6.22897}
6.8	7.02061	0.0735359	{6.85103, 7.19018}
8.1	7.98182	0.0817306	{7.79335, 8.17029}
9.3	8.94303	0.0960448	{8.72155, 9.16451}
9.9	9.90424	0.1142	{9.6409, 10.1676}
10.6	10.8655	0.134652	{10.5549, 11.176}

Ajust parabòlic

Versió amb excel

[1	x	x²]	y	y_c	$(y-y_c)^2$	Syc	y_{inf}	y_{sup}
1	1	1	8,70708	5,53145194	10,0846	10,8592	-20,1464	31,2093
1	2	4	31,5562	26,7670557	22,93582	7,2925	9,52306	44,0111
1	3	9	40,8246	53,1788182	152,6278	5,91371	39,1951	67,1625
1	4	16	86,1426	84,7667394	1,893053	6,10608	70,3282	99,2053
1	5	25	119,899	121,530819	2,663432	6,54031	106,065	136,996
1	6	36	162,986	163,471058	0,235104	6,54031	148,006	178,936
1	7	49	218,488	210,587455	62,41398	6,10608	196,149	225,026
1	8	64	248,792	262,880012	198,4855	5,91371	248,896	276,864
1	9	81	346,255	320,348727	671,1506	7,2925	303,105	337,593
1	10	100	368,406	382,9936	212,805	10,8592	357,316	408,671
				Sr^2	190,7564			

xT X	10	55	385
	55	385	3025
	385	3025	25333

t
2,36462
n
10

cov	263,88	-100,147	7,9481841
	-100,147	46,0272	-3,9740921
	7,94818	-3,97409	0,3612811

(xT X)-1	1,38333	-0,525	0,04167
	-0,525	0,24129	-0,02083
	0,04167	-0,02083	0,00189

		estimació	st error	{interval de confiança}	
$(xT\,X)^{-1}\,X^T\,Y=$	a	-10,528	16,2444	-48,9398	27,883781
	b	13,4714	6,78434	-2,57101	29,513742
	c	2,58808	0,60107	1,16679	4,0093735

xT Y	1632,06
	12436,4
	102261

Versió amb Mathematica

```
ClearAll["Global`*"]

llista = {{1, 8.707077353}, {2, 31.55619119}, {3, 40.82455458},
   {4, 86.14262193}, {5, 119.8988168}, {6, 162.986183},
   {7, 218.4877073}, {8, 248.7915225}, {9, 346.2553003},
   {10, 368.4057616}};

sol = LinearModelFit[llista, {1, x, x^2}, x]
```

FittedModel[$-10.528 + 13.4714\,x + 2.58808\,x^2$]

```
sol["CovarianceMatrix"] // MatrixForm
```

$$\begin{pmatrix} 263.88 & -100.147 & 7.94818 \\ -100.147 & 46.0272 & -3.97409 \\ 7.94818 & -3.97409 & 0.361281 \end{pmatrix}$$

```
sol["ParameterConfidenceIntervalTable"]
```

	Estimate	Standard Error	Confidence Interval
1	−10.528	16.2444	{−48.9398, 27.8838}
x	13.4714	6.78434	{−2.57104, 29.5138}
x^2	2.58808	0.601067	{1.16678, 4.00938}

```
sol["MeanPredictionConfidenceIntervalTable"]
```

Observed	Predicted	Standard Error	Confidence Interval
8.70708	5.53145	10.8592	{−20.1465, 31.2094}
31.5562	26.7671	7.2925	{9.52303, 44.0111}
40.8246	53.1788	5.91371	{39.1951, 67.1625}
86.1426	84.7667	6.10608	{70.3282, 99.2053}
119.899	121.531	6.54031	{106.065, 136.996}
162.986	163.471	6.54031	{148.006, 178.936}
218.488	210.587	6.10608	{196.149, 225.026}
248.792	262.88	5.91371	{248.896, 276.864}
346.255	320.349	7.2925	{303.105, 337.593}
368.406	382.994	10.8592	{357.316, 408.672}

Ajust no lineal

Se disposa de la taula adjunta de temperatures absolutes, pressions i volums d' un mol de gas real.
Se demana que es determinen (que s' ajusten) les constants a i b de Van der Waals d' aquest gas.
Recordeu que l' equació de Van der Waals és :

$$T = \frac{1}{R}\left(P + \frac{a}{v^2}\right)(v-b) \qquad \frac{\partial T}{\partial a} = \frac{1}{R}\frac{1}{v^2}(v-b) \qquad \frac{\partial T}{\partial b} = -\frac{1}{R}\left(P + \frac{a}{v^2}\right)$$

```
ClearAll["Global`*"];
tStudent = 2.01808;
RR = 0.082;
cap = {"V(L.)", "P(at.)", "T(K)"};

dades = {{1, 3.6, 100}, {1.2, 3.4, 100}, {1.4, 3.2, 100}, {1.6, 3, 100}, {1.8, 2.9, 100}, {2, 2.7, 100}, {2.2, 2.5, 100},
   {2.4, 2.4, 100}, {2.6, 2.3, 100}, {2.8, 2.2, 100}, {3, 2.1, 100}, {1, 7.9, 150}, {1.2, 7, 150}, {1.4, 6.3, 150},
   {1.6, 5.8, 150}, {1.8, 5.3, 150}, {2, 4.9, 150}, {2.2, 4.5, 150}, {2.4, 4.2, 150}, {2.6, 4, 150}, {2.8, 3.7, 150},
   {3, 3.5, 150}, {1, 12.6, 200}, {1.2, 10.9, 200}, {1.4, 9.6, 200}, {1.6, 8.6, 200}, {1.8, 7.8, 200}, {2, 7.1, 200},
   {2.2, 6.6, 200}, {2.4, 6.1, 200}, {2.6, 5.7, 200}, {2.8, 5.3, 200}, {3, 5, 200}, {1, 17.6, 250}, {1.2, 15, 250},
   {1.4, 13, 250}, {1.6, 11.5, 250}, {1.8, 10.4, 250}, {2, 9.4, 250}, {2.2, 8.6, 250}, {2.4, 7.9, 250}, {2.6, 7.4, 250},
   {2.8, 6.9, 250}, {3, 6.4, 250}};
n = Length[dades];

D2[a_, b_] = Sum[(dades[[i, 3]] - 1/RR (dades[[i, 2]] + a / dades[[i, 1]]^2) (dades[[i, 1]] - b))^2, {i, 1, n}];
eq1 = D[D2[a, b], a]; eq2 = D[D2[a, b], b];
sol = Solve[{eq1 == 0, eq2 == 0, a > 0, b > 0}, {a, b}];
Print[" "];
Print["a = ", sol[[1, 1, 2]], "    b = ", sol[[1, 2, 2]], "      sr^2 =", sr2 = D2[sol[[1, 1, 2]], sol[[1, 2, 2]]]/(n - 2)];
a = sol[[1, 1, 2]]; b = sol[[1, 2, 2]];
```

(* $Z^0_{iu} = \left(\frac{\partial f(x_u, \theta)}{\partial \theta_i}\right)_{\theta_0}$ *)

```
F = {};
For[i = 1, i ≤ n, i++,
  F = AppendTo[F, {1/RR 1/dades[[i, 1]]^2 (dades[[i, 1]] - b), -1/RR (dades[[i, 2]] + a/dades[[i, 1]]^2)}]
 ];
Print["Derivades = ", F];
```

(* $cov = (Z_0^T Z_0)^{-1} S_r^2$ *)

```
cov = Inverse[Transpose[F].F] sr2;
Print[" "];
Print["Variàncies-covariàncies dels paràmetres "]; Print["covariance matrix = ", cov // MatrixForm];
Print["a = ", a, "          b = ", b];
Print["s(a) = ", Sqrt[cov[[1, 1]]], ";     s(b) = ", Sqrt[cov[[2, 2]]]];
Print[" "];
```

(* $S^2_{y_i^c} = Z_i^T cov\, Z_i$; $Z_i = \begin{pmatrix} Z^0_{i1} & \ldots & Z^0_{ip} \end{pmatrix}$ *)

```
Print["Imprecissions de la y_calculada"];
s2yc = {};
For[i = 1, i ≤ n, i++,
  s2yc = AppendTo[s2yc, (F[[i]].cov.{{F[[i, 1]]}, {F[[i, 2]]}})[[1]]]];
taula = {{"y", "ycalc", "Syc", " ", "interval", "de", " confiança", " "}};
For[i = 1, i ≤ n, i++,
 taula = AppendTo[taula, {dades[[i, 3]], aux = 1/RR (dades[[i, 2]] + a / dades[[i, 1]]^2) (dades[[i, 1]] - b),
    aux2 = Sqrt[s2yc[[i]]], "{", aux - aux2 * tStudent, ",", aux + aux2 * tStudent, "}"}]
 ]; Print[Take[taula, 5] // TableForm];
Print["..."];
Print[Take[taula, -5] // TableForm];
```

```
a = 6.86314    b = 0.156852     sr^2 =3.40374

Derivades = {{10.2823, -127.599}, {8.83425, -99.5862}, {7.73487, -81.7268}, {6.87475, -69.2794}, {6.18469, -61.1982},
  {5.61935, -53.851}, {5.14803, -47.7805}, {4.74921, -43.799}, {4.40747, -40.43}, {4.11142, -37.5049}, {3.8525, -34.9094}, {10.2823, -180.038},
  {8.83425, -143.489}, {7.73487, -119.532}, {6.87475, -103.426}, {6.18469, -90.4665}, {5.61935, -80.6803}, {5.14803, -72.1708},
  {4.74921, -65.7502}, {4.40747, -61.1617}, {4.11142, -55.7976}, {3.8525, -51.9826}, {10.2823, -237.355}, {8.83425, -191.05},
  {7.73487, -159.776}, {6.87475, -137.572}, {6.18469, -120.954}, {5.61935, -107.51}, {5.14803, -97.7805}, {4.74921, -88.9209},
  {4.40747, -81.8934}, {4.11142, -75.3098}, {3.8525, -70.2753}, {10.2823, -298.331}, {8.83425, -241.05}, {7.73487, -201.239}, {6.87475, -172.938},
  {6.18469, -152.662}, {5.61935, -135.558}, {5.14803, -122.171}, {4.74921, -110.872}, {4.40747, -102.625}, {4.11142, -94.822}, {3.8525, -87.3484}}


Variàncies-covariàncies dels paràmetres

covariance matrix = ( 0.0191113     0.000948364  )
                    ( 0.000948364   0.0000520931 )

a = 6.86314          b = 0.156852

s(a) = 0.138244;     s(b) = 0.00721756


Imprecissions de la y_calculada
y      ycalc      Syc             interval    de     confiança
100    107.585    0.616588    {   106.341     ,      108.829      }
100    103.883    0.582639    {   102.707     ,      105.059      }
100    101.599    0.540675    {   100.507     ,      102.69       }
100    99.9805    0.4999      {   98.9716     ,      100.989      }
...
250    249.613    0.301815    {   249.004    ,    250.222    }
250    248.703    0.269602    {   248.159    ,    249.247    }
250    250.728    0.248938    {   250.226    ,    251.231    }
250    250.628    0.228009    {   250.168    ,    251.089    }
250    248.345    0.206966    {   247.927    ,    248.762    }
```

(* Fent ús de NonlinearModelFit *)

```
Clear[a, b]; solNLMF = NonlinearModelFit[dades, 1/RR (p + a / v^2) (v - b), {a, b}, {v, p}];
(* Print[solNLMF];*)
Print["Covariance =", solNLMF["CovarianceMatrix"] // MatrixForm, ";   ", "sr2 =", solNLMF["EstimatedVariance"], ";   ",
  solNLMF["ParameterConfidenceIntervalTable"]];
Print[ Take[solNLMF["MeanPredictionConfidenceIntervalTable"][[1, 1]], 5] // MatrixForm];
Print["..."];
Print[Drop[solNLMF["MeanPredictionConfidenceIntervalTable"][[1, 1]], {2, 40}] // MatrixForm];
```

Covariance = $\begin{pmatrix} 0.0191113 & 0.000948364 \\ 0.000948364 & 0.0000520931 \end{pmatrix}$; sr2 =3.40374;

	Estimate	Standard Error	Confidence Interval
a	6.86314	0.138244	{6.58415, 7.14212}
b	0.156852	0.00721756	{0.142286, 0.171417}

Observed	Predicted	Standard Error	Confidence Interval
100	107.585	0.616588	{106.341, 108.829}
100	103.883	0.582639	{102.707, 105.059}
100	101.599	0.540675	{100.507, 102.69}
100	99.9805	0.4999	{98.9716, 100.989}

...

Observed	Predicted	Standard Error	Confidence Interval
250	249.613	0.301815	{249.004, 250.222}
250	248.703	0.269602	{248.159, 249.247}
250	250.728	0.248938	{250.226, 251.231}
250	250.628	0.228009	{250.168, 251.089}
250	248.345	0.206966	{247.927, 248.762}

Ajust no lineal amb excel (Ajust de dades experimentals a l'equació de Van Deemter)

A	0,02480846
B	26,8048521
C	1,54867494

t	2,31
n	11

Eq. Teòrica
Y=A X+ B/X+C

$$cov = (Z_0^T Z_0)^{-1} S_r^2$$

$$S_{y_i^c}^2 = Z_i^T \, cov \, Z_i$$

$$Z_i = \begin{pmatrix} Z_{i1}^0 & \dots & Z_{ip}^0 \end{pmatrix}$$

X(mL/min)	Y(mm)	Y teorica	(Y-Yc)^2	∂Y/∂A = x	∂Y/∂B= 1/x	∂Y/∂C=1	Syc	y_{inf}	y_{sup}
3,4	9,59	9,52	0,01	3,4	0,294117647	1	0,11034715	9,26190182	9,77170565
7,1	5,29	5,50	0,04	7,1	0,14084507	1	0,05397389	5,37546661	5,62482599
16,1	3,63	3,61	0,00	16,1	0,062111801	1	0,05181047	3,49330665	3,73267101
20	3,42	3,39	0,00	20	0,05	1	0,05089314	3,26752363	3,50264995
23,1	3,46	3,28	0,03	23,1	0,043290043	1	0,04960845	3,16753811	3,39672914
34,4	3,06	3,18	0,01	34,4	0,029069767	1	0,04373915	3,08025942	3,28233431
40	3,25	3,21	0,00	40	0,025	1	0,04144147	3,11540495	3,30686453
44,7	3,31	3,26	0,00	44,7	0,022371365	1	0,04035411	3,16405635	3,35049232
65,9	3,5	3,59	0,01	65,9	0,015174507	1	0,04830205	3,4787253	3,70188076
78,9	3,86	3,85	0,00	78,9	0,012674271	1	0,06148399	3,70376658	3,98782261
96,8	4,24	4,23	0,00	96,8	0,010330579	1	0,08440633	4,03206512	4,42202236
		SUMA	0,110191648						
		Sr2	0,013773956						

Z^TZ	25974,5	11	430,4
	11	0,11704258	0,704985051
	430,4	0,70498505	11

$(Z^T Z)^{-1}$	0,00018844	0,04348822	-0,010160202
	0,04348822	23,9522	-3,236661149
	-0,0101602	-3,2366611	0,695886239

parametre	estimació	st error	{interval de confiança}	
A	0,02480846	0,00161107	0,021086899	0,02853003
B	26,8048521	0,57438363	25,4780259	28,1316783
C	1,54867494	0,09790356	1,322517721	1,77483215