Matemàtiques III:
0. Sistemes numèrics i fonts d'error

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1. Introducció
2. Codificacions numèriques
3. Conseqüències d'usar aritmètica de punt flotant

Continguts

1 Introducció

2 Codificacions numèriques

3 Conseqüències d'usar aritmètica de punt flotant

Errors d'arrodoniment i aritmètica computacional

Burden i Faires, sec. 1.2
Amat et al., cap. 1
Gilat i Subramaniam, cap. 1

¶¶ Els ordinadors representen qualsevol quantitat en binari i amb un nombre de xifres FINIT, mentre que entre els nombres reals hi ha alguns (irracionals) amb un nombre INFINIT de xifres.

Exemple: $\sqrt{3} \simeq 1.732050807568877$ (16 xifres significatives).

$\Longrightarrow$ L'operació $(\sqrt{3})^2 \simeq 2.999999999999999$ no donarà exactament 3.

$\Longrightarrow$ Es podrà representar de forma exacta els enters i els racionals amb un nombre de xifres menor o igual que el nombre de xifres significatives del nostre dispositiu (16 al cas de l'exemple). Exemples: $2+2=4$, $1/2+1/10=0.5+0.1=0.6$, $4.123 \times 5.1234 = 21.1237782$

¶ Els càlculs en un ordinador es realitzen amb representacions aproximades de nombres verdaders.$\Longrightarrow$ errors d'arrodoniment.

Continguts

1 Introducció

2 Codificacions numèriques

3 Conseqüències d'usar aritmètica de punt flotant

Representacions numèriques

¶¶ Els ordinadors emmagatzemen els nombres en forma de bits (dígits binaris) o bytes (paquets de 8 bits). Hi ha diferents tipus de representacions dels nombres que difereixen en la longitud de la paraula, en si es pren una representació de punt fix (enters) o punt flotant (reals), etc.

¶ Representació decimal (base 10): És l'habitual en els nostres càlculs (però no en la representació interna dels ordinadors). Exemple:

$$N = 2571 = 2 \cdot 10^3 + 5 \cdot 10^2 + 7 \cdot 10 + 1 \cdot 10^0.$$

Tot enter en base 10 té associat un polinomi ($p_N(x) = 2x^3 + 5x^2 + 7x + 1$ en l'exemple), de manera que $p_N(x = 10) = N$.

Notació:

$$N = (a_n a_{n-1} \dots a_0)_{10} = a_n 10^n + a_{n-1} 10^{n-1} + \dots + a_0 10^0, \quad 0 \leq a_i < 10.$$

¶ Representació binària (base 2): És la que fan servir els ordinadors, ja que té una equivalència directa a impuls elèctric (1) i la seua manca (0). Un enter no negatiu es representarà en binari com a:

$$N = (a'_n a'_{n-1} \dots a'_0)_2 = a'_n 2^n + a'_{n-1} 2^{n-1} + \dots + a'_0 2^0, \quad 0 \leq a'_i < 2.$$

N tindrà associat un polinomi $p'_N(x)$ tal que $p'_N(x = 2) = N$.

Conversió decimal-binari

¶ Enters: cal anar dividint per 2, fins que el quocient siga 0, fixant-se en la resta. LSB = Least Significant Digit. MSB = Most Significant digit.

Valor	Divisor	Resta	
43	2	1 ⟸ LSB	1×2^0
21	2	1	1×2^1
10	2	0	0×2^2
5	2	1	1×2^3
2	2	0	0×2^4
1	2	1 ⟸ MSB	1×2^5
0			

Podem comprovar que:

$$1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 43$$

$$(43)_{10} \equiv (101011)_2$$

¶ Reals: cal anar multiplicant per 2 la part fraccional i ens hem de fixar en la part entera (0 o 1). El procés acaba quan el resultat de la multiplicació siga $1.000\ldots$:

$0.3125 \times 2 = 0.625 \Longleftarrow$ MSB	0×2^{-1}
$0.625 \times 2 = 1.25$	1×2^{-2}
$0.25 \times 2 = 0.5$	0×2^{-3}
$0.5 \times 2 = 1.00 \Longleftarrow$ LSB	1×2^{-4}

Podem comprovar que:

$$0 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4} = 0.3125$$

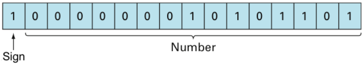
$$(0.3125)_{10} \equiv (0.0101)_2$$

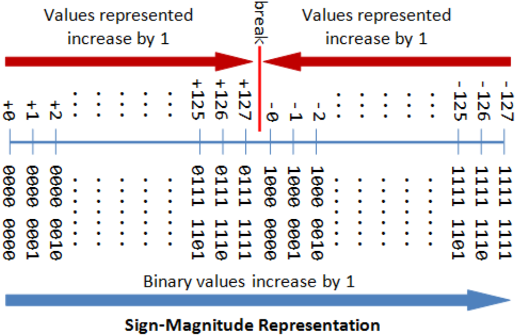
Exercici 1

Determineu la codificació binària dels següents nombres donats en format decimal: 725; 111.35; 0.1; 3.1415.

Codificacions dels enters amb signe: signe magnitud

Com codificaríem nombres enters negatius? En aquest cas sembla que la resposta òbvia en el cas binari és deixar un bit per a denotar el signe (positiu o negatiu). En el cas decimal és el que fem!; si no apareix res, és positiu, i si apareix el símbol $-$, és negatiu (això és un bit!!; 0/1). Arribem al que es coneix com a representació signe magnitud





Per a saber més:
http://www.ntu.edu.sg/home/ehchua/programming/java/DataRepresentation.html
http://www.unicode.org
Curiositats: no només es codifiquen els nombres sinó també els caràcters, i fins i tot els emojis!: http://www.unicode.org/reports/tr51/

Codificacions dels enters amb signe: complement a 2

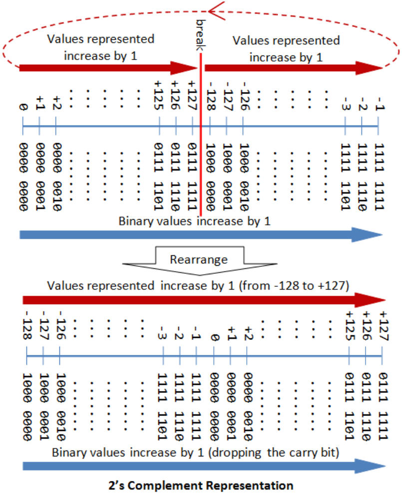
Signe magnitud presenta dos problemes:

a) el nombre 0 es codifica de dues formes;

b) es tracten els nombres positius i els negatius de forma separada.

Es planteja llavors la codificació complement a 2.

Amb aquesta codificació el nombre zero queda codificat d'una sola manera i es poden processar de la mateixa forma els nombres positius i negatius en les operacions d'addició/subtracció.



Per a pensar: Se t'acudeix alguna forma de codificació alternativa a la de complement a 2 que no tinga els problemes de la codificació signe magnitud?

Codificacions dels reals: coma fixa

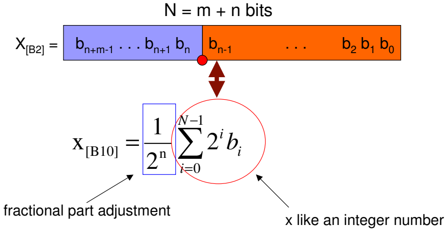
Com que tenim un nombre determinat de bits per a codificar els nombres (tant enters com reals), la primera possibilitat que se'ns acudeix és la codificació en coma fixa, en la qual prenem una part (m) dels bits disponibles (N) per a representar la part entera, i la resta ($n = N - m$) per a representar la part decimal de manera independent.

La coma fixa presenta el problema que quan multipliquem dos nombres en aquesta representació és molt habitual que el resultat *no càpia* en la codificació.

¶ Exemple:

Considera una codificació en coma fixa de 8 bits. Considerem $m = 4$ i $n = 4$. Sien els nombres 1234.5678 i 9876.5432. Quin és el resultat de la seua multiplicació? Pot representar-se en la codificació de coma fixa?

- $1234.5678 \times 9876.5432 = 12193262.210029$;
- Com que només tenim 4 bits davant de la coma en la representació de coma fixa considerada, el nombre 12193262.210029 desborda la representació (overflow).



$$x_{[B10]} = \frac{1}{2^n} \sum_{i=0}^{N-1} 2^i b_i$$

Per a pensar: Què passa quan multipliquem els nombres 0000.0200 i 0000.0001? Podries representar el resultat en el sistema de coma fixa l'exemple anterior? Hi ha en la bibliografia algun concepte en anglès que descriga el que succeeix en aquest cas?

Codificacions dels reals: punt flotant (I)

¶¶ Definició: Un sistema numèric de punt flotant és un subconjunt de nombres reals de la forma

$$y = \pm(0.d_1 d_2 \dots d_t)_\beta \times \beta^e = \pm(d_1\beta^{-1} + d_2\beta^{-2} + \dots + d_r\beta^{-t}) \cdot \beta^e \qquad (1)$$

- base: β.
- precisió o longitud de la paraula: t.
- rang dels exponents: $e_{\min} \leq e \leq e_{\max}$
- mantissa: $m = (0.d_1 d_2 \dots d_t)_\beta$, compleix $0 \leq m < 1$ i $0 \leq d_i < \beta$.
- Si el primer dígit significatiu compleix $d_1 \neq 0$, la representació (1) es diu que és normalitzada.
- Només hi ha una quantitat FINITA de nombres que podem representar amb punt flotant, sent el rang de nombres representables:

$$\beta^{e_{\min}-1} < y < (1 - \beta^{-t}) \cdot \beta^{e_{\max}}.$$

- Valors típics en Matlab (i en molts altres programes/compiladors amb precisió doble): $t = 53$, $\beta = 2$, $e_{\min} = -1021$, $e_{\max} = 1024$.

Codificacions dels reals: punt flotant (II)

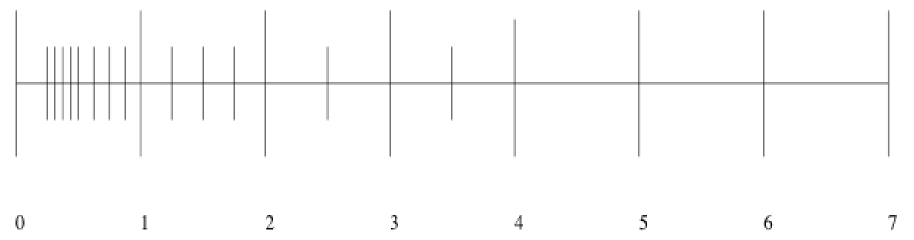
Exercici 2

Prenguem $\beta = 2$, $t = 3$, $e_{\min} = -1$ y $e_{\max} = 3$. Demostren que el sistema de punt flotant estaria format pels següents nombres no negatius (en base 10):

$$0, 0.25, 0.3125, 0.375, 0.4375, 0.5, 0.625, 0.75, 0.875,$$
$$1.0, 1.25, 1.50, 1.75, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0, 7.0,$$

Per què els valors 0.3125 i 0.4375 estan dins del sistema si la longitud de la paraula és $t = 3$?

Gràficament (noteu la densitat més alta de valors en l'interval $[0, 1]$):



Codificacions dels reals: punt flotant (III)

¶¶ Atès que en un sistema numèric de punt flotant normalitzat tenim un nombre finit de valors numèrics, com decidim el valor que fem servir per a representar qualsevol altre valor no comprès entre els elements d'aquest sistema? En l'exercici anterior, amb quin valor representaríem 0.46875? Noteu que $0.4375 < 0.46875 < 0.5$. Tenim 2 mètodes:

- Truncament: Tallem el nombre amb tants decimals com precisió t tinga el nostre sistema numèric. En el nostre cas, $t = 3$;
  $(0.46875)_{10} = (0.1111\ldots)_2 \times 2^{-1} \rightarrow (0.111)_2 \times 2^{-1} = (0.4375)_{10}$, per tant, la representació de punt flotant en el nostre sistema seria $fl(0.46875) = 0.4375$.
  Genèricament trunquem la mantissa:

$$\pm 0.d_1 d_2 \cdots d_t d_{t+1} d_{t+2} \cdots \simeq \pm 0.d_1 d_2 \cdots d_t$$

- Arrodoniment: Es prenen tantes xifres com determina la longitud de la paraula i en l'última es manté igual si la següent xifra és menor que la base dividida per dos i se li suma una unitat (a l'última xifra) si la següent és més gran que cinc. En el cas anterior, el resultat de l'arrodoniment seria:
  $(0.46875)_{10} = (0.1111\ldots)_2 \times 2^{-1} \rightarrow (0.100)_2 \times 2^{0} = (0.5)_{10}$ per tant $fl(0.4478) = 0.5$. Genèricament:

$$\begin{aligned} \pm 0.d_1 d_2 \ldots d_t d_{t+1} d_{t+2} \ldots &\simeq \pm 0.d_1 d_2 \cdots d_t & \text{si } d_{t+1} < \beta/2 \\ \pm 0.d_1 d_2 \ldots d_t d_{t+1} d_{t+2} \ldots &\simeq \pm 0.d_1 d_2 \cdots (d_t + 1) & \text{si } d_{t+1} > \beta/2 \end{aligned}$$

Continguts

1 Introducció

2 Codificacions numèriques

3 Conseqüències d'usar aritmètica de punt flotant

Cancel·lació catastròfica

- Exemple: $p = 0.54617$ i $q = 0.54601$. El valor exacte de $r = p - q$ és $r = 0.00016$. Calculem la resta amb una aritmètica que només tinga 4 xifres significatives:
  - Amb arrodoniment: $p^* \equiv fl(p) = 0.5462$ i $q^* \equiv fl(q) = 0.5460$, de manera que $r^* \equiv fl(r) = fl(p^* - q^*) = 0.0002$, la qual és una aproximació de 4 xifres de r, però vegeu que

    $$\text{error relatiu} \equiv \frac{|r - r^*|}{|r|} = \frac{|0.00016 - 0.0002|}{|0.00016|} = 0.25,$$

    és a dir, que tenim un 25% d'error relatiu, i el resultat només té una xifra significativa!.
  - Amb truncament: $p^* = 0.5461$, $q^* = 0.5460$, $r^* = 0.0001$ i, per tant,

    $$\frac{|r - r^*|}{|r|} = \frac{|0.00016 - 0.0001|}{|0.00016|} = 0.375,$$

    en aquest cas, l'error relatiu és major, encara que el resultat també té només una xifra significativa.

IMPORTANT

En restar dues quantitats molt semblants es poden perdre moltes xifres significatives.

Insensibilitat del resultat operant amb nombres molt dispars

- Exemple: Prenem en un sistema de punt flotant amb $t = 3$ i $\beta = 10$ dos nombres $p = (0.518) \cdot 10^3$ i $q = (0.437) \cdot 10^0$. Si sumem tots dos nombres, hauríem de tenir $r = p + q =$ 518.437, però a causa de l'aritmètica de punt flotant tenim

$$r* = fl(p+q) = fl(518 + .437) = fl(518.437) = (.518) \cdot 10^3 = p^*!.$$

Noteu, però, que l'error relatiu d'aquest càlcul és

$$\frac{|r - r^*|}{|r|} = \frac{|518.437 - 518|}{|518.437|} = 0.00084,$$

IMPORTANT

L'element neutre de la suma NO és únic en aritmètica de punt flotant.

No compliment de la propietat associativa

- En general $fl([fl(a) + fl(b)] + fl(c)) \neq fl(fl(a) + [fl(b) + fl(c)])$.
- Exemple: Prenem tres nombres $a = 63.1$, $b = 4.24$ i $c = 0.247$, en un sistema de punt flotant amb $t = 3$ i $\beta = 10$ de manera que $a^* = (.631) \cdot 10^2$ $b^* = (.424) \cdot 10^1$, i $c = (.247) \cdot 10^0$. A causa de l'aritmètica de punt flotant tenim

$$
\begin{aligned}
(a^* + b^*) + c^* &= ((.631) \cdot 10^2 + (.424) \cdot 10^1) + (.247) \cdot 10^0 \\
&= (.673) \cdot 10^2 + (.247) \cdot 10^0 = (.675) \cdot 10^2, \\
a^* + (b^* + c^*) &= (.631) \cdot 10^2 + ((.424) \cdot 10^1 + (.247) \cdot 10^0) \\
&= (.631) \cdot 10^2 + (.448) \cdot 10^1 = (.676) \cdot 10^2 .
\end{aligned}
$$

Noteu que l'error relatiu d'aquest càlcul és $\dfrac{|67.577 - 67.5|}{|67.577|} = 0.0011.$

IMPORTANT

L'ordre de les operacions altera el resultat.

En Matlab, fent `x=(1+1e20)-1e20` dóna 0,
però si fem `x=1+(1e20-1e20)` dóna 1!

Dues operacions inverses no retornen el resultat original

- En el cas de la multiplicació i la divisió,

$$fl[fl(a) \cdot fl(1/a)] \neq 1.$$

Exemple: Prenem $a = 3$ en aquest cas, $1/a = 0,333333333\ldots$. En un sistema de punt flotant amb $t = 3$ i $\beta = 10$ de manera que $a^* = (.300) \cdot 10^1$, i $(1/a)^* = (.333) \cdot 10^0$ tenim

$$a^* \cdot (1/a)^* = (.999) \cdot 10^0 \neq 1\,.$$

- En el cas de la suma i la resta,

$$fl[(fl(a) + fl(b)) - fl(b)] \neq f(a),$$

però

$$fl[fl(a) + (fl(b) - fl(b))] = f(a)\,.$$

Exemple Matlab:

```
matlab> format long
matlab> a=4/3
a =  1.33333333333333
matlab> b=a-1
b =  0.333333333333333
matlab> c=3*b
c =  1.000000000000000
matlab> e=1-c
e =  2.22044604925031e-16
```

IMPORTANT

Com a conseqüència de la pèrdua de la propietat associativa en aritmètica de punt flotant, operacions inverses no NECESSÀRIAMENT retornen el valor original.

Conclusions

Si bé és cert que els errors d'arrodoniment són normalment petits, quan es repeteixen multitud d'ocasions en el context d'algoritmes numèrics molt llargs i complexos, poden donar lloc a efectes catastròfics:

- L'explosió del coet Arianne el 4 de juny de 1996, generada per un *overflow* a l'ordinador de bord. Es va perdre el coet i la seua càrrega, amb un valor estimat de 500 milions de dòlars. (Més informació a `http://www.around.com/ariane.html`).

- El fracàs de la missió d'un míssil americà *patriot* durant la guerra del Golf el 1991 (va causar la mort de 28 soldats), a causa d'un error d'arrodoniment en el càlcul de la seua trajectòria.

- En l'operació de llastrament controlat de la base de formigó de la plataforma d'extracció de petroli Sleipner A (Noruega, 23 d'agost de 1991) va sorgir una fuita i es va enfonsar de forma catastròfica. El pes total de la plataforma era d'unes 97.000 tones, de manera que va caure 220 m fins al llit marí i va originar un terratrèmol d'intensitat 3.0 en l'escala de Richter, així com una pèrdua de 700 milions de dòlars. L'error es va deure, en part, a l'aproximació inexacta dels elements finits del model per al càlcul de l'estructura (utilitzant el popular programa NASTRAN). Els esforços de cisallament se subestimaren en un 47%, de manera que es van produir parets de formigó que no eren prou gruixudes.

Matemàtiques III:
1. Regressió lineal

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1. Introducció
2. Recta de regressió
3. El coeficient de correlació mostral
4. Ajustaments relacionats amb la recta de regressió
5. Regressió robusta
6. Regressió múltiple

Continguts

1 Introducció

2 Recta de regressió

3 El coeficient de correlació mostral

4 Ajustaments relacionats amb la recta de regressió

5 Regressió robusta

6 Regressió múltiple

Per a què serveix la regressió?

Donat un conjunt de variables (p. ex., mesures d'intensitat d'un senyal radioelèctric), l'ús tècnic d'aquestes requereix conèixer com es relacionen entre elles i quines són les variables més importants.

En la pràctica volem expressar la relació entre diferents variables en forma matemàtica (p. ex., en forma d'equació). Quan el coneixement d'una variable

a. determina completament el valor d'una altra $\Longrightarrow$ relació exacta o funcional.
b. no permet inferir cap informació sobre una altra $\Longrightarrow$ variables independents.
c. es pot predir, amb més o menys precisió, el valor de l'altra $\Longrightarrow$ relació estocàstica o estadística.

¶¶ El cas c és el més habitual en aplicacions cientificotècniques, estadística, etc., i necessitem instruments matemàtics (regressió, inferència estadística, etc.) per a quantificar com és la relació funcional subjacent entre diferents variables. Aquests instruments matemàtics han de permetre *predir* el valor de les variables dependents, trobar *anomalies* (p. ex., degudes a errors en les mesures) i estimar el nivell de confiança d'aquesta inferència.

Continguts

1. Introducció
2. Recta de regressió
3. El coeficient de correlació mostral
4. Ajustaments relacionats amb la recta de regressió
5. Regressió robusta
6. Regressió múltiple

Ajust de dades

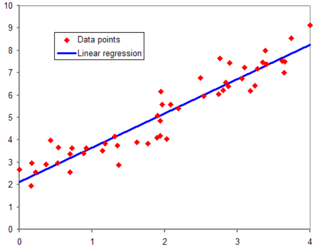
Bibliografia: Colomer, Cap. 17, Ref. b2.

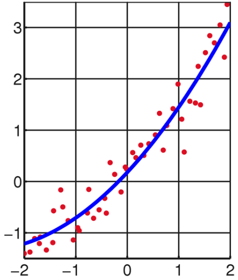
¶¶ **Mètodes de regressió**: Estudi de la relació entre variables numèriques

Exemples: Relació entre alçada i pes de xiquets nounats, relació entre el voltatge, la intensitat i la impedància ($V = Z \cdot I$), problemes de calibratge (en experiments de laboratori)

Diagrames de dispersió i corbes d'ajust

Hi ha infinites relacions entre variables, però podem establir dos grans grups: les lineals i les no lineals. Les primeres són aquelles en què les relacions entre les variables s'estableixen mitjançant una sèrie de coeficients constants; en les segones apareixen funcions no lineals en la relació entre variables. En el que segueix els w_k són coeficients constants (nombres) i les y i x_i variables:





Lineals

$$y = w_0 + w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_n \cdot x_n$$

No lineals

$$y = w_0 + w_1 \cdot (x_1)^2 + w_2 \cdot \log x_2 + \cdots + w_n \cdot (x_n)^{-1}$$

Ajust de dades

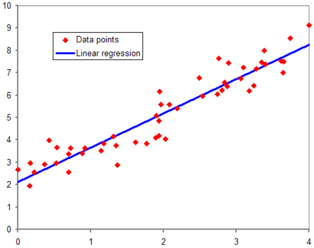
Bibliografia: Colomer, Cap. 17, Ref. b2.

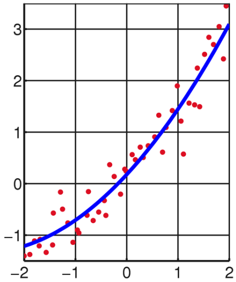
¶¶ **Mètodes de regressió**: Estudi de la relació entre variables numèriques

Exemples: Relació entre alçada i pes de xiquets nounats, relació entre el voltatge, la intensitat i la impedància ($V = Z \cdot I$), problemes de calibratge (en experiments de laboratori)

Diagrames de dispersió i corbes d'ajust

Hi ha infinites relacions entre variables, però podem establir dos grans grups: les lineals i les no lineals. Les primeres són aquelles en què les relacions entre les variables s'estableixen mitjançant una sèrie de coeficients constants; en les segones apareixen funcions no lineals en la relació entre variables. En el que segueix els w_k són coeficients constants (nombres) i les y i x_i variables:





Lineals

$$y = w_0 + w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_n \cdot x_n$$

En aquest curs només veurem aquestes així com la forma de determinar els paràmetres w_k i les mesures d'error que s'utilitzen per a saber si la relació és correcta o no.

No lineals

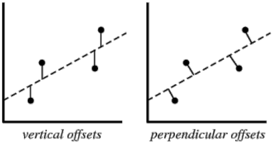
$$y = w_0 + w_1 \cdot (x_1)^2 + w_2 \cdot \log x_2 + \cdots + w_n \cdot (x_n)^{-1}$$

Regressió lineal simple (I)

Bibliografia: Colomer, Cap. 17, Ref. b2.

¶¶ La **recta de regressió** o **recta de mínims quadrats** $y = a + bx$, *és la que minimitza els quadrats de les desviacions* (ϵ_i), és a dir, la funció:

$$\Phi(a,b) = \sum_{i}^{n} \underbrace{(y_i - (a + bx_i))}_{\epsilon_i}^{2}$$



Carl Friedrich Gauss

¶¶ **Equacions normals** per a determinar a i b $\Longrightarrow$ Condicions de mínim per a $\Phi(a,b)$:

$$\frac{\partial \Phi}{\partial a} = 0 \quad (1)$$

$$\frac{\partial \Phi}{\partial b} = 0 \quad (2)$$

¶¶ **Paràmetres de la recta de regressió**: Solució del sistema anterior $\left(\bar{x} := \frac{1}{n}\sum_{i}^{n} x_i,\ \bar{y} := \frac{1}{n}\sum_{i}^{n} y_i\right)$

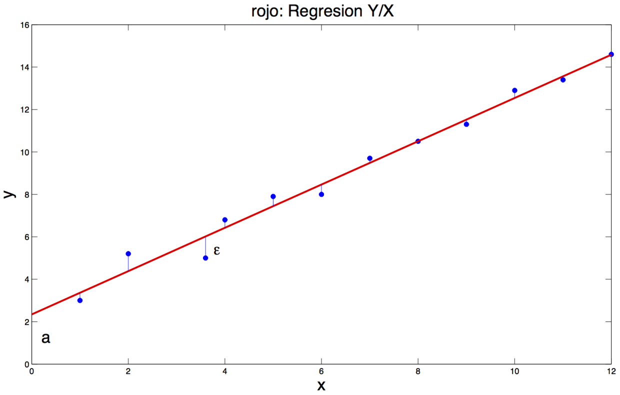
$$b = \frac{s_{xy}}{s_{xx}} \quad , \quad a = \bar{y} - b\bar{x} = \bar{y} - \frac{s_{xy}}{s_{xx}}\bar{x} \quad (3)$$

- $s_{xy} = \frac{1}{n}\sum_{i}^{n}(x_i - \bar{x})(y_i - \bar{y})$
- $s_{xx} = \frac{1}{n}\sum_{i}^{n}(x_i - \bar{x})^2$
- $s_{yy} = \frac{1}{n}\sum_{i}^{n}(y_i - \bar{y})^2$

s_{xx}, s_{xy}, i s_{yy} són les variàncies mostrals.

Regressió lineal simple (II): Interpretació geomètrica dels paràmetres

$$\Phi(a,b) = \sum_i^n (\underbrace{y_i - (a + bx_i)}_{\epsilon_i})^2$$



¶¶ **Paràmetres de la recta de regressió**: Solució del sistema anterior

$$b = \frac{s_{xy}}{s_{xx}} \quad , \quad a = \bar{y} - b\bar{x} = \bar{y} - \frac{s_{xy}}{s_{xx}}\bar{x} \tag{4}$$

- $s_{xy} = \frac{1}{n}\sum_i^n (x_i - \bar{x})(y_i - \bar{y})$
- $s_{xx} = \frac{1}{n}\sum_i^n (x_i - \bar{x})^2$
- $s_{yy} = \frac{1}{n}\sum_i^n (y_i - \bar{y})^2$

s_{xx}, s_{xy}, i s_{yy} són les variàncies mostrals.

Regressió lineal simple (III): Deducció de (4)

De (1) i (2) resulta: $\Longrightarrow$

$$\frac{\partial \Phi}{\partial a} = -2 \sum_{i}^{n} [y_i - (a + b\, x_i)] = 0 \tag{5}$$

$$\frac{\partial \Phi}{\partial b} = -2 \sum_{i}^{n} [y_i - (a + b\, x_i)]\; x_i = 0 \tag{6}$$

De (5)

$$\sum_{i}^{n} y_i - n\, a - b \sum_{i}^{n} x_i = 0 \Longrightarrow a = \bar{y} - b\, \bar{x} \tag{7}$$

De (6)

$$\sum_{i}^{n} y_i\, x_i - a \sum_{i}^{n} x_i - b \sum_{i}^{n} x_i\, x_i = 0 \Longrightarrow b = \frac{\sum_i^n (y_i - \bar{y})\, x_i}{\sum_i^n (x_i - \bar{x})\, x_i} \tag{8}$$

Tenint en compte que es compleix:

$$\sum_{i}^{n} (y_i - \bar{y})\, \bar{x} = 0 \quad , \quad \sum_{i}^{n} (x_i - \bar{x})\, \bar{x} = 0 \tag{9}$$

podem reescriure la solució per b com en (4):

$$b = \frac{\sum_i^n (y_i - \bar{y})\, x_i}{\sum_i^n (x_i - \bar{x})\, x_i} = \frac{\left(\sum_i^n (y_i - \bar{y})\, x_i\right) - \left(\sum_i^n (y_i - \bar{y})\, \bar{x}\right)}{\left(\sum_i^n (x_i - \bar{x})\, x_i\right) - \left(\sum_i^n (x_i - \bar{x})\, \bar{x}\right)} = \frac{s_{xy}}{s_{xx}} \tag{10}$$

Regressió lineal simple (IV): Relacions addicionals

$$
\begin{aligned}
s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) &= \frac{1}{n}\left(\sum_{i=1}^{n}x_iy_i - \sum_{i=1}^{n}y_i\overline{x} - \sum_{i=1}^{n}x_i\overline{y} + \sum_{i=1}^{n}\overline{x}\cdot\overline{y}\right)\\
&= \frac{1}{n}\sum_{i=1}^{n}x_iy_i - \overline{x}\underbrace{\frac{1}{n}\sum_{i=1}^{n}y_i}_{\equiv\overline{y}} - \overline{y}\underbrace{\frac{1}{n}\sum_{i=1}^{n}x_i}_{\equiv\overline{x}} + \overline{x}\cdot\overline{y}\frac{1}{n}\underbrace{\sum_{i=1}^{n}1}_{\equiv n}\\
&= \frac{1}{n}\sum_{i=1}^{n}x_iy_i - \overline{x}\cdot\overline{y} - \overline{y}\cdot\overline{x} + \overline{x}\cdot\overline{y} = \frac{1}{n}\sum_{i=1}^{n}x_iy_i - \overline{x}\cdot\overline{y}
\end{aligned}
$$

Anàlogament es pot demostrar (Exercici 1):

$$
s_{xx} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2 \qquad s_{yy} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2 = \frac{1}{n}\sum_{i=1}^{n}y_i^2 - \overline{y}^2
$$

Les quantitats s_{xx}, s_{yy}, s_{xy} són les components de la matriu de covariàncies de X i Y, definides com a

$$
\mathrm{cov}(X,Y) = \begin{bmatrix} s_{xx} & s_{xy} \\ s_{xy} & s_{yy} \end{bmatrix}
$$

En Matlab la matriu de covariàncies es pot obtenir utilitzant l'ordre `cov(x,y,1)`.

Regressió lineal simple (V): Implementació en Matlab

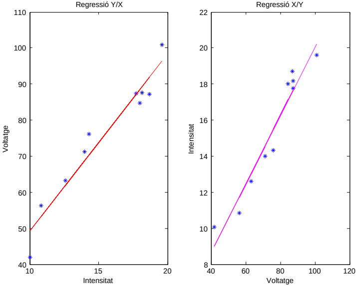
Tensió (V)	Intensitat (mA)
100.90	19.62
42.02	10.05
87.33	17.75
87.57	18.17
87.13	18.69
56.25	10.84
71.12	14.00
63.13	12.60
84.67	18.00
76.08	14.31

Introduïm les dades I (és la X) i V (és la Y), en sengles VECTORS de Matlab:

```
» I = [19.62, 10.05, 17.75, 18.17, 18.69, 10.84, 14., 12.6, 18., 14.31];
» V = [100.9, 42.02, 87.33, 87.53, 87.13, 56.25, 71.12, 63.13, 84.67, 76.08];
```

Valors de la recta de regressió de V sobre I: $Vr_i = a + bI_i = \bar{V} + b(I_i - \bar{I})$, utilitzant la fórmula (4). Anàlogament s'obté $Ir_i = \alpha + \beta V_i = \bar{I} + \beta(V_i - \bar{V})$, per a la recta de regressió de I sobre V, per a la qual tenim $I = \alpha + \beta V$, on $\alpha = \overline{I} - \beta\overline{V}$ i $\beta = s_{IV}/s_{VV}$.

```
» C=cov(I,V,1); % matriu covariància
» b=C(1,2)/ C(1,1); % b
» bet=C(1,2)/ C(2,2); % beta
» Vr=mean(V)+b*(I-mean(I)) % valors de V-Reg-I
» Ir=mean(I)+bet*(V-mean(V)) % valors de I-Reg-V
» plot(I,V,'*',I,Vr,'r')
» title('roig:  Regressió Y/X')
» plot(V,I,'*',V,Ir,'m')
» title('magenta Regressió X/Y')
```



Amb el càlcul anterior obtenim els següents valors:

$s_{II} = \texttt{C(1,1)} = 10.89$, $\quad s_{IV} = \texttt{C(1,2)} = 53.75$,

$\bar{V} = \texttt{mean(V)} = 75.62$, $\quad \bar{I} = \texttt{mean(I)} = 15.40$,

$$V = 4.93 \cdot I - 0.35, \qquad I = 0.19 \cdot V + 0.96$$

Segons l'ajust la resistència del dispositiu és de 4.93 kiloohms. El terme independent (també conegut com a biaix) fa referència als errors de mesura.

Regressió lineal simple (VI): resum

y = a + b x

La recta de regressió de $Y = (y_1, \ldots, y_n)$ sobre $X = (x_1, \ldots, x_n)$ és $y = a + b\, x$ amb

$$a = \bar{y} - b\bar{x} \quad , \quad b = \frac{s_{xy}}{s_{xx}}$$

$$y - \bar{y} = \frac{s_{xy}}{s_{xx}}(x - \bar{x})$$

x = c + d y

La recta de regressió de $X = (x_1, \ldots, x_n)$ sobre $Y = (y_1, \ldots, y_n)$ és $x = c + d\, y$ amb

$$c = \bar{x} - d\bar{y} \quad , \quad d = \frac{s_{xy}}{s_{yy}}$$

$$x - \bar{x} = \frac{s_{xy}}{s_{yy}}(y - \bar{y})$$

IMPORTANT

La recta de regressió de $Y = (y_1, \ldots, y_n)$ sobre $X = (x_1, \ldots, x_n)$ deixa bé el mateix nombre de punts per sobre i per sota en el diagrama de dispersió si n és parell, bé el mateix nombre de punts més/menys un si n és senar.

Continguts

1. Introducció
2. Recta de regressió
3. El coeficient de correlació mostral
4. Ajustaments relacionats amb la recta de regressió
5. Regressió robusta
6. Regressió múltiple

El coeficient de correlació mostral (I)

¶¶ **Definició:**

$$r \equiv \frac{s_{xy}}{\sqrt{s_{xx}}\sqrt{s_{yy}}}$$

¶¶ **Propietats:**

- **P1**: La recta de regressió de $Y = (y_1, \ldots, y_n)$ sobre $X = (x_1, \ldots, x_n)$ es pot escriure com a:

$$\frac{y - \bar{y}}{\sqrt{s_{yy}}} = r\frac{x - \bar{x}}{\sqrt{s_{xx}}}$$

- **P2**: La recta de regressió de $X = (x_1, \ldots, x_n)$ sobre $Y = (y_1, \ldots, y_n)$, es pot escriure

$$\frac{x - \bar{x}}{\sqrt{s_{xx}}} = r\frac{y - \bar{y}}{\sqrt{s_{yy}}}$$

- **P3**: Els pendents de les dues rectes tenen el mateix signe, i són diferents (excepció: $r = \pm 1$).
- **P4**: El coefficient de correlació està *normalitzat*: $-1 \leq r \leq 1$
- **P5**: Com més a prop està r^2 d'1 menor és la dispersió pel que fa a les rectes de regressió, i més forta és la lligadura lineal entre els valors de les mostres X i Y .
  Prova: Cal observar que

$$D_y \equiv \frac{1}{n}\sum_{i=1}^{n}(y_i - (a + bx_i))^2 = s_{yy}(1 - r^2)$$

Anàlogament, $$D_x \equiv \frac{1}{n}\sum_{i=1}^{n}(x_i - (\alpha + \beta y_i))^2 = s_{xx}(1 - r^2)$$

Considerant que $D_x = s_{xx}(1 - r^2)$ i $D_y = s_{yy}(1 - r^2)$, resulta evident que per a $r = 0$ les desviacions (tant en x com a y) són màximes, i quan $r = \pm 1$, són mínimes.

El coeficient de correlació mostral (II)

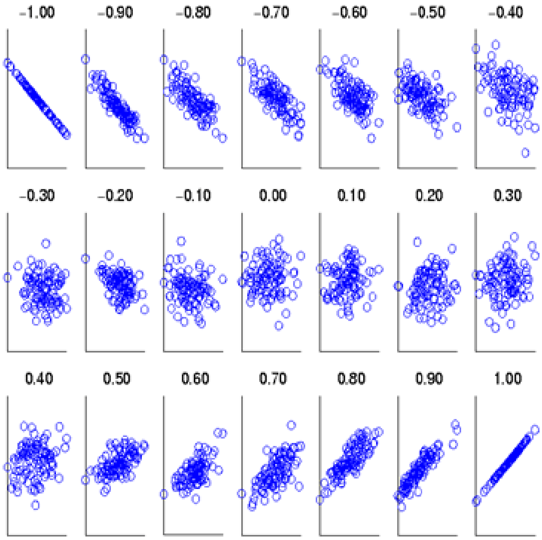


Figure 5.11. Scatter plots illustrating correlations from -1 to 1.

IMPORTANT

- És un índex molt útil, ja que dóna idea de la possible relació lineal, del seu tipus i *direcció*.
- Si $r = 0$ no vol dir que no hi haja relació entre les variables sinó que no hi ha relació lineal entre elles.

Continguts

1 Introducció

2 Recta de regressió

3 El coeficient de correlació mostral

4 Ajustaments relacionats amb la recta de regressió

5 Regressió robusta

6 Regressió múltiple

Ajustaments relacionats amb la recta de regressió (I)

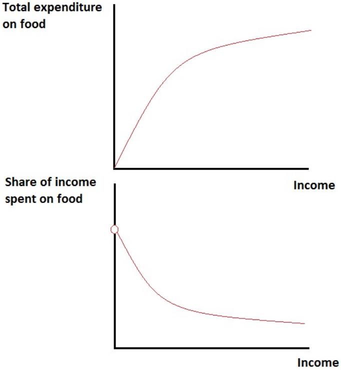
¶¶ Siga el conjunt de dades: $Y = (y_1, y_2, \ldots, y_n)$ $X = (x_1, x_2, \ldots, x_n)$.

- **Ajust hiperbòlic**: Es tracta d'ajustar per una funció de la forma

$$y = a + b\frac{1}{x}$$

En aquest cas es poden calcular els paràmetres a i b (*desconeguts a priori*) calculant la recta de regressió per a les dades:

$$Y = (y_1, y_2, \ldots, y_n)\,,$$
$$Z = (\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}).$$



La corba d'Engel que expressa la demanda d'un bé en funció de la renda, adopta de vegades la forma d'una hipèrbola equilàtera.

Ajustaments relacionats amb la recta de regressió (II)

- **Ajust potencial**: Ajustar a una corba de la forma

$$y = ax^b$$

Tenint en compte que $\log y = \log a + b \log x$, es poden calcular els paràmetres determinant la recta de regressió per les dades:

$$Y = (\log y_1, \log y_2, \ldots, \log y_n) \quad , \quad X = (\log x_1, \log x_2, \ldots, \log x_n).$$

Exemple: La topologia d'Internet es pot quantificar en termes de lleis de potències. El famós article de Faloutsos et al. (1999) proporciona les lleis que regeixen la configuració i el creixement d'internet en termes de lleis de potències que s'obtenen ajustant dades reals.

Power-Law 1 (rank exponent) *The outdegree, d_v, of a node v, is proportional to the rank of the node, r_v, to the power of a constant, $\mathcal{R}$:*

$$d_v \propto r_v^{\mathcal{R}}$$

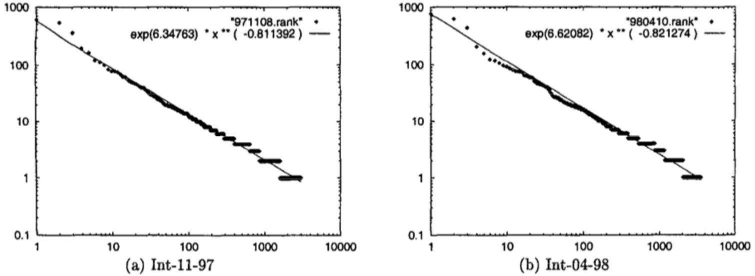


Figure 3: The rank plots. Log-log plot of the outdegree d_v versus the rank r_v in the sequence of decreasing outdegree.

Ajustaments relacionats amb la recta de regressió (III)

- Ajust exponencial: Ajustar a una corba de la forma

$$y = ab^x$$

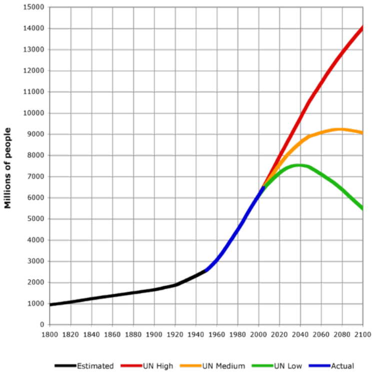
Tenint en compte que

$$\log y = \log a + x \log b,$$

es poden calcular els paràmetres calculant la recta de regressió per a les dades

$$\begin{aligned} Y &= (\log y_1, \log y_2, \ldots, \log y_n) \\ X &= (x_1, x_2, \ldots, x_n). \end{aligned}$$

Models matemàtics usuals per al creixement d'una població: l'exponencial (Malthus 1798) i el logístic (Verhulst 1838).



Exemple: S'ha mesurat la concentració d'una substància en una reacció obtenint els següents resultats:

T (seg.)	0	2	4	6	8	10
mols / litre	10.0	6.8	4.9	4.1	3.5	3.1

Obtenir l'expressió de la concentració en funció del temps mitjançant

- regressió lineal
- ajust exponencial

Exemple

¶¶ Siguen les dades: `x=[ 0 2 4 6 8 10]` ; `y=[10.0 6.8 4.9 4.1 3.5 3.1]`

Recta de regressió de Y sobre X:

$$y - \bar{y} = \frac{s_{xy}}{s_{xx}}(x - \bar{x}), \quad \equiv \quad y = a + bx, \quad \begin{cases} b = \frac{s_{xy}}{s_{xx}} \\ a = \bar{y} - b\bar{x} \end{cases}$$

$$s_{xy} = -7.53, \qquad s_{xx} = 11.67 \qquad \bar{x} = 5, \qquad \bar{y} = 5.4$$

Resultat: $y - 5.4 = -0.6457(x - 5), \quad \equiv \quad y = 8.6286 - 0.6457x$

Ajust exponencial: $y = a \cdot b^x$

$$\log(y) = \log(a) + x\log(b) \qquad \equiv \qquad z = A + Bx$$

amb $A = \log(a)$, $B = \log(b)$

Calculem la recta de regressió per $z = \log(y)$, respecte a x, i obtenim
$A = 2.1740$, $B = -0.1147 \implies a = e^A = 8.7932, \qquad b = e^B = 0.8917$
La funció de l'ajust és doncs: $y = 8.7932 \cdot 0.8917^x$

Ajustaments relacionats amb la recta de regressió (IV)

- Ajust logístic: Ajustar a una corba de la forma

$$y = \frac{1}{1+e^{-(a+bx)}}$$

Tenint en compte que

$$1-y = \frac{e^{-(a+bx)}}{1+e^{-(a+bx)}}$$

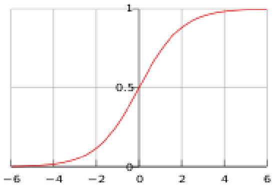
resulta que

$$\log\frac{y}{1-y} = \log e^{a+bx} = a+bx,$$

es poden calcular els paràmetres calculant la recta de regressió per a les dades

$$Y = \left(\log\frac{y_1}{1-y_1}, \log\frac{y_2}{1-y_2}, \dots, \log\frac{y_n}{1-y_n}\right)$$
$$X = (x_1, x_2, \dots, x_n).$$

La funció logística pren valors entre 0 i 1, i pot utilitzar-se com una probabilitat per a classificar esdeveniments.



La regressió logística és un model molt utilitzat per a classificació de patrons (per exemple per a reconeixement automàtic de rutes, diagnòstic, xarxes neuronals, etc.).

Continguts

1 Introducció

2 Recta de regressió

3 El coeficient de correlació mostral

4 Ajustaments relacionats amb la recta de regressió

5 Regressió robusta

6 Regressió múltiple

Valors atípics (*outliers*)

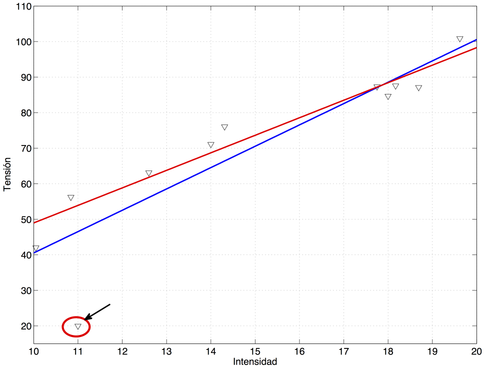
En l'obtenció de la resistència a partir de les mesures de voltatge en funció de la intensitat, què passaria a l'ajust si féssem una mesura *errònia*?

Tensió (V)	Intensitat (mA)
100.90	19.62
42.02	10.05
87.33	17.75
87.57	18.17
87.13	18.69
56.25	10.84
71.12	14.00
63.13	12.60
84.67	18.00
76.08	14.31
20	**11**

Si fem servir totes les dades: $V = 6 \cdot I - 19.5$

Si no fem servir la mesura dolenta: $V = 4.93 \cdot I - 0.35$

Representem les dades (roig: ajust sense l'últim valor, i blau: ajust amb totes les dades)



Veiem que la dada atípica (outlier) fa *baixar* la recta per a reduir l'error corresponent a aquest punt.

Regressió robusta

El problema dels outliers rau en el fet que s'intenta minimitzar la suma dels residus al quadrat, cosa que fa que tinguen molta importància en l'ajust final.

Possibles solucions:

a) Definir una nova funció a minimitzar. (*funció de cost robusta*)

b) *Pesar* cadascun dels termes que apareixen a la funció a minimitzar.

c) *Eliminar* aquells valors que presenten un valor alt d'error.

Regressió robusta

El problema dels outliers rau en el fet que s'intenta minimitzar la suma dels residus al quadrat, cosa que fa que tinguen molta importància en l'ajust final.

Possibles solucions:

a) Definir una nova funció a minimitzar. (*funció de cost robusta*)

b) *Pesar* cadascun dels termes que apareixen a la funció a minimitzar.

c) *Eliminar* aquells valors que presenten un valor alt d'error.

En l'ajust per mínims quadrats minimitzem: $\phi(a,b) = \sum_i^n (\epsilon_i)^2$, on $\epsilon_i = y_i - (a + b \cdot x_i)$, és a dir, cada error *es pesa* amb el seu quadrat.

Alternativa (1): Donem *menys pes* a cada error canviant la funció a minimitzar:

$$\phi(a,b) = \sum_i^n \log\left(\cosh\left(\epsilon_i\right)\right)$$

$$\phi(a,b) = \sum_i^n |\epsilon_i|$$

El problema de totes és que no hi ha una solució directa (com la quadràtica) i cal aplicar procediments iteratius. Ho veurem en pròxims temes!

Alternativa (2): *Pesem* cada error amb un factor w_k multiplicant cada terme d'error:

$$\phi(a,b) = \sum_i^n w_k (\epsilon_i)^2$$

El factor w_k pot tenir en compte l'error comès en ajustar el model, per això pot caldre plantejar un mètode iteratiu per a trobar el mínim d'aquesta funció.

Continguts

1 Introducció

2 Recta de regressió

3 El coeficient de correlació mostral

4 Ajustaments relacionats amb la recta de regressió

5 Regressió robusta

6 Regressió múltiple

Regressió múltiple

El cas de relacionar un parell de variables no és el més comú en enginyeria; més aviat ens trobarem amb la situació de relacionar un gran nombre de variables entre si. Per exemple la sortida d'un filtre digital no deixa de ser una relació lineal entre diferents instants temporals del senyal d'entrada al filtre.

En aquest cas tindrem la següent expressió:

$$y^k = a_0 + a_1 \cdot x_1^k + a_2 \cdot x_2^k + \cdots + a_m \cdot x_m^k$$

Si definim:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_m^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_m^2 \\ & \cdots & & & \\ 1 & x_1^n & x_2^n & \cdots & x_m^n \end{bmatrix}; \mathbf{Y} = \begin{bmatrix} y^1 \\ y^2 \\ \cdots \\ y^n \end{bmatrix}; \mathbf{A} = \begin{bmatrix} a_0 \\ a_1 \\ \cdots \\ a_m \end{bmatrix}$$

tindrem: $\mathbf{Y} = \mathbf{X} \cdot \mathbf{A}$

Es pot plantejar el mateix procediment que en la regressió simple: minimitzar la suma dels errors al quadrat.

$$\phi(a,b) = \sum_k^n (\epsilon_k)^2 \Rightarrow \phi(\mathbf{A}) = \sum_k^n \left(y_i^k - (a_0 + a_1 \cdot x_1^k + a_2 \cdot x_2^k + \cdots + a_m \cdot x_m^k) \right)^2$$

Vectorialment: $\phi(\mathbf{A}) = (\mathbf{Y} - \mathbf{X} \cdot \mathbf{A})^T \cdot (\mathbf{Y} - \mathbf{X} \cdot \mathbf{A}) = \mathbf{Y}^T \cdot \mathbf{Y} - 2 \cdot \mathbf{Y}^T \cdot \mathbf{X} \cdot \mathbf{A} + \mathbf{A}^T \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{A}$

Derivant respecte als paràmetres i igualant a zero arribem a les **equacions normals**:

$$\mathbf{A} = [\mathbf{X}^T \cdot \mathbf{X}]^{-1} \cdot \mathbf{X}^T \cdot \mathbf{Y}$$

Matemàtiques III:
2. Mètodes numèrics per a la resolució de sistemes d'equacions lineals

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

Resolució algebraica (I): Introducció

Objectiu

Siguen $\mathbf{A}$ *una matriu* $n \times n$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ *sengles vectors columna de dimensió* n. *Es tracta de trobar la solució* $\mathbf{x}$ *del sistema de* n *equacions lineals:* $\mathbf{A}\mathbf{x} = \mathbf{b} \;\Rightarrow\; \sum_{j=1}^{n} a_{ij}\, x_j \;=\; b_i \;,\; i = 1 \ldots n$

¶ *Un sistema d'equacions lineals (sistema lineal d'equacions, o sistema lineal)* és un conjunt d'equacions lineals com ara

$$\left\{ \begin{array}{rcrcrcr} 6\,x_1 & + & 4\,x_2 & + & 2\,x_3 & = & 2 \\ x_1 & + & x_2 & + & 2\,x_3 & = & -1 \\ -2\,x_1 & + & x_2 & - & 2\,x_3 & = & 0 \end{array} \right. \tag{1}$$

¶ Tipus de sistemes:

I.- Incompatible $\Longrightarrow$ No té cap solució ($|A| = 0$, $\mathbf{b} \neq \mathbf{0}$).
II.- Compatible $\Longrightarrow$ Té alguna solució.
II.a.- Compatible determinat $\Longrightarrow$ Té una única solució ($|A| \neq 0$).
II.b.- Compatible indeterminat $\Longrightarrow$ Té un conjunt infinit de solucions ($|A| = 0, \mathbf{b} = \mathbf{0}$).

Esquemàticament:

$$Tipus\ de\ sistemes \begin{cases} Compatible \begin{cases} Determinat \\ Indeterminat \end{cases} \\ Incompatible \end{cases} \tag{2}$$

Resolució algebraica (II): Substitució i reducció

Si els càlculs foren realitzats sense arrodoniment, ens conduirien a la solució exacta del sistema.

Substitució

¶ *Consisteix a aïllar, en una de les equacions, qualsevol incògnita i substituir-la en les altres equacions.*

¶ Exemple:

$$\begin{cases} x & +y & = & 2 \\ x & -y & = & 0 \end{cases} \tag{3}$$

1. A la segona equació, seleccionem, p. ex., la incògnita x i l'aïllem: $x = y$
2. Substituïm en la primera equació la incògnita x pel valor anterior:
   $y + y = 2 \quad \Rightarrow 2y = 2 \quad \Rightarrow y = 1$
3. El resultat $y = 1$ el portem a alguna de les dues equacions originals (p. ex. la primera):
   $x + 1 = 2 \quad \Rightarrow x = 1$

Reducció

¶ *Consisteix –per a sistemes de dues equacions– a transformar de manera que n'obtinguem dues més en què una mateixa incògnita aparega amb el mateix coeficient i signe diferent. En sumar es cancel·la aquesta.*

¶ Exemple:

$$\begin{cases} x & +y & = & 2 \\ x & -y & = & 0 \end{cases} \tag{4}$$

En l'exemple anterior, simplement sumant ambdues equacions cancel·lem la incògnita y:
$2x = 2 \quad \Rightarrow x = 1 \quad (\ldots)$

Resolució algebraica (III): Mètode de Cramer

Siga el sistema de n equacions lineals:

$$\mathbf{Ax} = \mathbf{b} \;\Leftrightarrow\; \sum_{j=1}^{n} a_{ij}\; x_j \;=\; b_i \;,\; (i, j =, 1 \ldots, n) \tag{5}$$

on $\mathbf{A}$ és una matriu de coeficients reals, no singular, de dimensió $n \times n$, i $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ sengles vectors columna de dimensió n.

Solució general:

$$\mathbf{x} \;=\; \mathbf{A}^{-1}\,\mathbf{b} \tag{6}$$

on $\mathbf{A}^{-1}$ és la matriu inversa de $\mathbf{A}$

¶¶ **Algorisme:**

$$x_i \;=\; \frac{det\ \mathbf{B_i}}{det\ \mathbf{A}} \;,\; (i = 1, \ldots, n) \tag{7}$$

on la matriu $\mathbf{B_i}$ és

$$\begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix} \tag{8}$$

¶ **Cost computacional:** Nombre d'operacions: $\mathcal{N} \approx 2(n+1)!$

Resolució algebraica (III): Mètode de Cramer

Siga el sistema de n equacions lineals:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \;\Leftrightarrow\; \sum_{j=1}^{n} a_{ij}\, x_j \;=\; b_i \;,\; (i, j =, 1 \ldots, n) \tag{5}$$

on $\mathbf{A}$ és una matriu de coeficients reals, no singular, de dimensió $n \times n$, i $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ sengles vectors columna de dimensió n.

Solució general:

$$\mathbf{x} \;=\; \mathbf{A}^{-1}\,\mathbf{b} \tag{6}$$

on $\mathbf{A}^{-1}$ és la matriu inversa de $\mathbf{A}$

¶¶ **Algorisme:**

$$x_i \;=\; \frac{det\ \mathbf{B_i}}{det\ \mathbf{A}} \;,\; (i = 1, \ldots, n) \tag{7}$$

on la matriu $\mathbf{B_i}$ és

$$\begin{pmatrix} a_{11} & \ldots & a_{1,i-1} & b_1 & a_{1,i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,i-1} & b_2 & a_{2,i+1} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,i-1} & b_n & a_{n,i+1} & \ldots & a_{nn} \end{pmatrix} \tag{8}$$

¶ **Cost computacional:** Nombre d'operacions: $\mathcal{N} \approx 2(n+1)! \Longrightarrow$ per a un PC típic de 10 gigaflops:

- $n = 16 \;\Longrightarrow T_{\text{calcul}} \approx (2 \cdot 17!)/10^{10} = 711374856192000/10^{10} \simeq 7.1 \cdot 10^4$ s ≈ 20 h

Resolució algebraica (III): Mètode de Cramer

Siga el sistema de n equacions lineals:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \;\Leftrightarrow\; \sum_{j=1}^{n} a_{ij}\, x_j \;=\; b_i \;,\; (i, j =, 1 \ldots, n) \tag{5}$$

on $\mathbf{A}$ és una matriu de coeficients reals, no singular, de dimensió $n \times n$, i $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ sengles vectors columna de dimensió n.

Solució general:

$$\mathbf{x} \;=\; \mathbf{A}^{-1}\, \mathbf{b} \tag{6}$$

on $\mathbf{A}^{-1}$ és la matriu inversa de $\mathbf{A}$

¶¶ **Algorisme:**

$$x_i \;=\; \frac{det\ \mathbf{B_i}}{det\ \mathbf{A}} \;,\; (i = 1, \ldots, n) \tag{7}$$

on la matriu $\mathbf{B_i}$ és

$$\begin{pmatrix} a_{11} & \ldots & a_{1,i-1} & b_1 & a_{1,i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,i-1} & b_2 & a_{2,i+1} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \cdots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,i-1} & b_n & a_{n,i+1} & \ldots & a_{nn} \end{pmatrix} \tag{8}$$

¶ **Cost computacional:** Nombre d'operacions: $\mathcal{N} \approx 2(n+1)! \Longrightarrow$ per a un PC típic de 10 gigaflops:

- $n = 16 \;\Longrightarrow T_{\text{calcul}} \approx (2 \cdot 17!)/10^{10} = 711374856192000/10^{10} \simeq 7.1 \cdot 10^4$ s ≈ 20 h
- $n = 20 \;\Longrightarrow T_{\text{calcul}} \approx 324$ anys

Resolució algebraica (III): Mètode de Cramer

Siga el sistema de n equacions lineals:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \;\Leftrightarrow\; \sum_{j=1}^{n} a_{ij}\; x_j \;=\; b_i \;,\; (i,j =, 1\ldots, n) \tag{5}$$

on $\mathbf{A}$ és una matriu de coeficients reals, no singular, de dimensió $n \times n$, i $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ sengles vectors columna de dimensió n.

Solució general:

$$\mathbf{x} \;=\; \mathbf{A}^{-1}\;\mathbf{b} \tag{6}$$

on $\mathbf{A}^{-1}$ és la matriu inversa de $\mathbf{A}$

¶¶ **Algorisme:**

$$x_i \;=\; \frac{det\;\mathbf{B_i}}{det\;\mathbf{A}} \;,\; (i = 1, \ldots, n) \tag{7}$$

on la matriu $\mathbf{B_i}$ és

$$\begin{pmatrix} a_{11} & \ldots & a_{1,i-1} & b_1 & a_{1,i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,i-1} & b_2 & a_{2,i+1} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,i-1} & b_n & a_{n,i+1} & \ldots & a_{nn} \end{pmatrix} \tag{8}$$

¶ **Cost computacional:** Nombre d'operacions: $\mathcal{N} \approx 2(n+1)! \Longrightarrow$ per a un PC típic de 10 gigaflops:

- $n = 16 \;\Longrightarrow\; T_{\mathrm{calcul}} \approx (2 \cdot 17!)/10^{10} = 711374856192000/10^{10} \simeq 7.1 \cdot 10^4$ s ≈ 20 h
- $n = 20 \;\Longrightarrow\; T_{\mathrm{calcul}} \approx 324$ anys
- $n = 100 \Longrightarrow T_{\mathrm{calcul}} \approx 10^{142}$ anys

Resolució algebraica (IV): Sistemes amb solució immediata

La matriu $\mathbf{A}$ és diagonal: $\mathbf{A} = \mathbf{D}$

¶ Exemple

$$\begin{cases} 2x_1 &= 1 \\ 3x_2 &= 4 \\ x_3 &= 6 \end{cases} \Rightarrow \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 4/3 \\ 6 \end{pmatrix}$$

¶¶ **Algorisme:**

$$x_i = \frac{b_i}{d_{ii}} \quad , \quad (i = 1, \dots, n)$$

on els d_{ii} són els termes de la diagonal.

La matriu $\mathbf{A}$ és triangular superior: $\mathbf{A} = \mathbf{U}$

¶¶ **Algorisme:**

$$x_n = \frac{b_n}{u_{nn}} \quad , \quad x_i = \frac{1}{u_{ii}} \left(b_i - \sum_{j=i+1}^{n} u_{ij}\, x_j \right) \quad , \quad (i = n-1, n-2, \dots, 1) \tag{9}$$

on els u_{ij} són els termes de la matriu triangular superior.

La matriu $\mathbf{A}$ és triangular inferior: $\mathbf{A} = \mathbf{L}$

¶¶ **Algorisme:**

$$x_1 = \frac{b_1}{l_{11}} \quad , \quad x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{j=1}^{i-1} l_{ij}\, x_j \right) \quad , \quad (i = 2, \dots, n) \tag{10}$$

on els l_{ij} són els termes de la matriu triangular inferior.

¶ **Cost computacional:** Nombre d'operacions elementals: $\mathcal{N} = n$ (primer cas), $\mathcal{N} \approx n^2$ en els altres.

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

Mètodes directes: Gauss (I)

¶¶ *Consisteix a triangular la matriu augmentada del sistema mitjançant transformacions elementals, fins a obtenir equacions d'una sola incògnita.*

¶ Exemple

$$\left\{\begin{array}{l}(-1/2)x+(-1/2)y+(-1/2)z=-15\\ 2x+6y=4z+40\\ 3x+3y=6z\end{array}\right. \longrightarrow \left\{\begin{array}{cccccc}x & +y & +z & = & 30\\ x & +3y & -2z & = & 20\\ x & +y & -2z & = & 0\end{array}\right. \tag{11}$$

1. Restant la primera equació a les dues següents:

$$\left\{\begin{array}{ccccc}x & +y & +z & = & 30\\ & 2y & -3z & = & -10\\ & & -3z & = & -30\end{array}\right. \tag{12}$$

2. En aquest cas, en la tercera equació s'han eliminat la x i la y $\quad \Rightarrow z=10$

3. Substituint z en la segona equació $\quad \Rightarrow y=10$

4. Substituint z i y en la primera equació $\quad \Rightarrow x=10$

Mètodes directes: Eliminació Gaussiana (II)

¶¶ Siga el sistema de n equacions lineals:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \;\Rightarrow\; \sum_{j=1}^{n} a_{ij}\, x_j \;=\; b_i \;,\; i = 1 \ldots n \tag{13}$$

on $\mathbf{A}$ és una matriu $n \times n$ de coeficients reals, no singular, de dimensió $n \times n$, i $\mathbf{x}, \mathbf{b} \in \Re^n$ sengles vectors columna de dimensió n. Redefinint $b_i \;=:\; a_{i\,,\,n+1}$, el sistema (13) és:

$$\left\{ \begin{array}{ccccccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & + & \ldots & + & a_{1n}\,x_n & = & a_{1\,,\,n+1} \\ a_{21}\,x_1 & + & a_{22}\,x_2 & + & \ldots & + & a_{2n}\,x_n & = & a_{2\,,\,n+1} \\ \ldots & + & \ldots & + & \ldots & + & \ldots & = & \ldots \\ a_{n1}\,x_1 & + & a_{n2}\,x_2 & + & \ldots & + & a_{nn}\,x_n & = & a_{n\,,\,n+1} \end{array} \right. \tag{14}$$

¶ **Eliminació gaussiana:**

1. Suposem que $a_{11} \neq 0 \;\Rightarrow$ En (14) restem el múltiple $\dfrac{a_{i1}}{a_{11}}$ de la primera equació a la $i-$èsima equació $(i = 2, \ldots n) \;\Rightarrow$

$$\left\{ \begin{array}{ccccccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & + & \ldots & + & a_{1n}\,x_n & = & a_{1\,,\,n+1} \\ & & a_{22}^{(1)}\,x_2 & + & \ldots & + & a_{2n}^{(1)}\,x_n & = & a_{2\,,\,n+1}^{(1)} \\ & & \ldots & & \ldots & & \ldots & = & \ldots \\ & & a_{n2}^{(1)}\,x_2 & + & \ldots & + & a_{nn}^{(1)}\,x_n & = & a_{n\,,\,n+1}^{(1)} \end{array} \right. \tag{15}$$

Els nous coeficients $a_{ij}^{(1)}$ són donats per

$$a_{ij}^{(1)} \;=\; a_{ij} \;-\; \left(\frac{a_{i1}}{a_{11}}\right) a_{1j} \;\;,\;\; (i = 2, \ldots, n) \;\;,\;\; (j = 2, \ldots, n+1) \tag{16}$$

Mètodes directes: Eliminació Gaussiana (III)

2. Suposem que $a_{22}^{(1)} \neq 0 \quad \Rightarrow$ En (15) restem el múltiple $\dfrac{a_{i2}^{(1)}}{a_{22}^{(1)}}$ de la segona equació a la $i-$èsima equació $(i = 3, \ldots n) \quad \Rightarrow$

$$\left\{\begin{array}{ccccccccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & + & a_{13}\,x_3 & + & \ldots & + & a_{1n}\,x_n & = & a_{1\,,\,n+1} \\ & & a_{22}^{(1)}\,x_2 & + & a_{23}^{(1)}\,x_3 & + & \ldots & + & a_{2n}^{(1)}\,x_n & = & a_{2\,,\,n+1}^{(1)} \\ & & & & a_{33}^{(2)}\,x_3 & + & \ldots & + & a_{3n}^{(2)}\,x_n & = & a_{3\,,\,n+1}^{(2)} \\ & & & & \ldots & & \ldots & & \ldots & = & \ldots \\ & & & & a_{n3}^{(2)}\,x_3 & & \ldots & + & a_{nn}^{(2)}\,x_n & = & a_{n\,,\,n+1}^{(2)} \end{array}\right. \tag{17}$$

Els nous coeficients $a_{ij}^{(2)}$ són donats per

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \left(\frac{a_{i2}^{(1)}}{a_{22}^{(1)}}\right) a_{2j}^{(1)} \quad , \quad (i = 3, \ldots, n) \quad , \quad (j = 3, \ldots, n+1) \tag{18}$$

3. Continuant aquest procés arribem al <u>sistema final</u>:

$$\left\{\begin{array}{ccccccccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & + & a_{13}\,x_3 & + & \ldots & + & a_{1n}\,x_n & = & a_{1\,,\,n+1} \\ & & a_{22}^{(1)}\,x_2 & + & a_{23}^{(1)}\,x_3 & + & \ldots & + & a_{2n}^{(1)}\,x_n & = & a_{2\,,\,n+1}^{(1)} \\ & & & & a_{33}^{(2)}\,x_3 & + & \ldots & + & a_{3n}^{(2)}\,x_n & = & a_{3\,,\,n+1}^{(2)} \\ & & & & \ldots & & \ldots & & \ldots & = & \ldots \\ & & & & & & \ldots & & \ldots & = & \ldots \\ & & & & & & & & \ldots & = & \ldots \\ & & & & & & & & a_{nn}^{(n-1)}\,x_n & = & a_{n\,,\,n+1}^{(n-1)} \end{array}\right. \tag{19}$$

Mètodes directes: Eliminació gaussiana (IV)

En (19) els coeficients $a_{ij}^{(k)}$ són donats per

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \left(\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \right) a_{kj}^{(k-1)} \qquad (20)$$

on els índexs satisfan:

$k = 1, \ldots, n-1 \quad ; \quad j = k+1, \ldots, n+1 \quad ; \quad i = k+1, \ldots, n \quad ; \quad a_{ij}^{(0)} := a_{ij}$

4. Solució de (13)

$$\boxed{x_i = \left(\frac{1}{a_{ii}^{(i-1)}} \right) \left(a_{i,n+1}^{(i-1)} - \sum_{j=i+1}^{n} a_{ij}^{(i-1)} \, x_j \right) \quad , \quad i = n, \ldots, 1 \qquad (21)}$$

¶ **Cost computacional:** El nombre d'operacions: $\mathcal{N} \approx n^3$ per a un PC típic de 10 gigaflops:

- $n = 16 \implies T_{\text{calcul}} \approx 16^3/10^{10} = 4096/10^{10} \approx 4 \cdot 10^{-7}$ s
- $n = 20 \implies T_{\text{calcul}} \approx 8 \cdot 10^{-7}$ s
- $n = 100 \implies T_{\text{calcul}} \approx 10^{-4}$ s

Eliminació Gaussiana (V): Exemple

Exemple 1: En (13, 14) fem $n = 2$, i apliquem l'eliminació gaussiana

Solució:

1. El sistema (14) es redueix a:

$$\left\{ \begin{array}{ccccc} a_{11}\, x_1 & + & a_{12}\, x_2 & = & a_{1\,,\,3} \\ a_{21}\, x_1 & + & a_{22}\, x_2 & = & a_{2\,,\,3} \end{array} \right. \tag{22}$$

2. Suposem que $a_{11} \neq 0 \;\; \Rightarrow$ En (22) restem el múltiple $\dfrac{a_{21}}{a_{11}}$ de la primera equació a la 2a equació $\Rightarrow$

$$\left\{ \begin{array}{ccccc} a_{11}\, x_1 & + & a_{12}\, x_2 & = & a_{1\,,\,3} \\ \underbrace{\left[a_{21} - \left(\dfrac{a_{21}}{a_{11}}\right) a_{11}\right]}\, x_1 & + & \underbrace{\left[a_{22} - \left(\dfrac{a_{21}}{a_{11}}\right) a_{12}\right]}\, x_2 & = & \underbrace{a_{2\,,\,3} - \left(\dfrac{a_{21}}{a_{11}}\right) a_{1\,,\,3}} \\ \mathbf{0} \quad x_1 & + & \mathbf{a_{22}^{(1)}} \quad x_2 & = & \mathbf{a_{23}^{(1)}} \end{array} \right.$$

Els nous coeficients $a_{ij}^{(1)}$ són, en aquest cas, $a_{2j}^{(1)}$, $j = 2, 3$ i són donats per

$$a_{2j}^{(1)} = a_{2j} - \left(\frac{a_{21}}{a_{11}}\right) a_{1j} \quad , \quad (j = 2, 3) \tag{23}$$

$$a_{22}^{(1)} = a_{22} - \left(\frac{a_{21}}{a_{11}}\right) a_{12} \tag{24}$$

$$a_{23}^{(1)} = a_{23} - \left(\frac{a_{21}}{a_{11}}\right) a_{13} \tag{25}$$

Eliminació Gaussiana (VI): Anàlisi de l'error (I)

¶ Com que en un ordinador treballem amb un sistema de punt flotant, els errors d'arrodoniment i de truncament fan que ni tan sols els valors que introduïm com a elements de la matriu $\mathbf{A}$ o el vector $\mathbf{b}$ siguen *exactes*, sinó **APROXIMATS**.

¶ Per tant, quan calculem la solució del sistema $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ usant un mètode d'eliminació gaussiana, que hauria de donar un resultat *exacte*, ens trobarem que la solució serà només **APROXIMADA**.

¶ Per a avaluar la bondat de l'aproximació obtinguda, podem fer servir el nombre de condició, $\kappa(\mathbf{A})$ definit com a

$$\kappa(\mathbf{A}) \;=\; ||\mathbf{A}||_F \cdot ||\mathbf{A}^{-1}||_F$$

on $||\mathbf{A}||_F$ representa la norma de $\mathbf{A}$ en el sentit euclidià:

$$||\mathbf{A}||_F = \left(\sum_{i,j=1}^{n} a_{ij}^2 \right)^{1/2}$$

IMPORTANT

Si $\kappa(\mathbf{A}) \gg 1$ diem que la matriu $\mathbf{A}$ **està mal condicionada**. Si això passa, petits canvis en el sistema d'equacions poden produir grans canvis en la solució.

Eliminació Gaussiana (VI): Anàlisi de l'error (II) [Matriu mal condicionada]

¶ Exemple

Considereu la matriu

$$\mathbf{A} = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}, \qquad \text{amb inversa:} \qquad \mathbf{A}^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$$

la qual té un nombre de condició $\kappa(\mathbf{A}) \gg 1$ (COMPROVAR!):

$$\kappa(\mathbf{A}) = ||\mathbf{A}||_2 \cdot ||\mathbf{A}^{-1}||_2 = \sqrt{7^2 + 10^2 + 5^2 + 7^2} \cdot \sqrt{(-7)^2 + 10^2 + 5^2 + (-7)^2} = 223$$

Si resolem els següents sistemes, veurem que la solució és molt diferent (tot i que només canvia lleugerament el terme independent).

$$\begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1.00 \\ 0.70 \end{pmatrix} \implies \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.00 \\ 0.10 \end{pmatrix}$$

$$\begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1.01 \\ 0.69 \end{pmatrix} \implies \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -0.17 \\ 0.22 \end{pmatrix}$$

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

Eliminació gaussiana i descomposició $LU(I)$

¶¶ *El procés d'eliminació gaussiana es pot interpretar com la descomposició de la matriu de coeficients del sistema en un producte de dues matrius*

$$A = LU$$

on L és triangular inferior i U triangular superior.

Exemple 1 Considerem el següent sistema 2×2, i la seua corresponent notació matricial,

$$\left\{ \begin{array}{rcl} 3x+y & = & 2 \\ x-y & = & 0 \end{array} \right\}; \quad A = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Realitzem el procés d'eliminació gaussiana, i escrivim la notació en forma matricial:

Eliminem la incògnita x de la segona equació.

- (nova equació 2) = -3 x (equació 2) + (equació 1)

$$\left\{ \begin{array}{rcl} 3x+y & = & 2 \\ x-y & = & 0 \end{array} \right\}; \quad \equiv \quad \left\{ \begin{array}{rcl} 3x+y & = & 2 \\ 4y & = & 2 \end{array} \right\};$$

$$AX = b \qquad \equiv \qquad A^{(1)}X = b^{(1)}$$

$$\begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \equiv \qquad \begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

- El nou sistema és equivalent a l'anterior, ja que la nova equació és una **combinació lineal** de les equacions del sistema.
- La matriu de coeficients del nou sistema és triangular superior.

Eliminació Gaussiana i Descomposició $LU(II)$

Per a entendre la relació que hi ha entre les matrius A i $A^{(1)}$, cal observar que l'operació

- (nova equació 2) = -3 x (equació 2) + (equació 1)

és el resultat d'encadenar 2 transformacions elementals:

1. Multiplicar la segona equació per -3
2. Sumar la primera i la segona equacions, i deixar el resultat en la segona equació.

Com que les equacions del sistema estan associades a les files de la matriu A, aquestes operacions es tradueixen en actuacions similars sobre les files de la matriu A, que es poden efectuar fàcilment pre-multiplicant per certes **matrius elementals**.

1. Multiplicar la segona fila de A per -3

$$\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -3 & 3 \end{pmatrix}$$

2. Sumar la primera i la segona files, deixant el resultat en la segona

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ -3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix} = A^{(1)}$$

Eliminació Gaussiana i Descomposició $LU(III)$

Cal observar que l'operació

- (nova equació 2) = -3 x (equació 2) + (equació 1)

és el resultat d'encadenar dues transformacions elementals que es poden expressar en termes de productes de matrius:

1. Multiplicar la segona equació per -3:

$$\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ -3\,a_{21} & -3\,a_{22} \end{pmatrix}$$

2. Sumar la primera i la segona equacions, i deixar el resultat en la segona equació:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ -3\,a_{21} & -3\,a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{11} - 3\,a_{21} & a_{12} - 3\,a_{22} \end{pmatrix}$$

En l'exemple anterior:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix} = A^{(1)} \;=:\; U$$

$\Longrightarrow$

$$\begin{pmatrix} 1 & 0 \\ 1 & -3 \end{pmatrix} A = U$$

$\Longrightarrow$

$$A = \begin{pmatrix} 1 & 0 \\ 1 & -3 \end{pmatrix}^{-1} U = \begin{pmatrix} 1 & 0 \\ 1/3 & -1/3 \end{pmatrix} U = L\,U$$

Eliminació Gaussiana i Descomposició $LU(IV)$

Per tant, podem escriure

$$AX = b \quad \equiv \quad \begin{pmatrix} 1 & 0 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} X = \begin{pmatrix} 1 & 0 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \equiv$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} X = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \equiv \quad A^{(1)}X = b^{(1)}$$

Així, el procés d'eliminació gaussiana sobre el sistema es pot interpretar com un procés de factorització, en el qual trobem una matriu triangular inferior que permet passar de la forma original a la forma triangular superior.

En l'exemple anterior, s'ha eliminat la variable multiplicant per -3 la segona equació i sumant, i és evident que hi ha llibertat sobre el tipus de combinacions que es poden fer amb les equacions per a eliminar aquesta variable.

No obstant això, des del punt de vista de **l'organització algorísmica**, és important seguir un **criteri**.

Eliminació Gaussiana i Descomposició $LU(V)$

El criteri que se segueix en els paquets informàtics, traduït a aquest exemple, és el de

- multiplicar la primera equació per $-1/3$ i sumar-la a la segona.

Aquesta operació s'aconsegueix multiplicant per la matriu

$$M_1 = \begin{pmatrix} 1 & 0 \\ -1/3 & 1 \end{pmatrix}$$

Així, podem escriure

$$AX = b \equiv \begin{pmatrix} 1 & 0 \\ -1/3 & 1 \end{pmatrix} A\ X = \begin{pmatrix} 1 & 0 \\ -1/3 & 1 \end{pmatrix} b \quad \equiv \quad \begin{pmatrix} 3 & 1 \\ 0 & -4/3 \end{pmatrix} X = \begin{pmatrix} 2 \\ -2/3 \end{pmatrix} \equiv \tilde{A}^{(1)} X = \tilde{b}^{(1)}$$

La matriu M_1 és una matriu triangular inferior amb uns en la diagonal. M_1 és una matriu invertible, i és fàcil comprovar que

$$M_1^{-1} = \begin{pmatrix} 1 & 0 \\ 1/3 & 1 \end{pmatrix}$$

Observeu que M_1^{-1} és també una matriu triangular inferior amb uns a la diagonal, és a dir, és del mateix tipus que M_1. És molt fàcil comprovar que per a calcular la inversa d'aquest tipus de matrius només cal canviar de signe l'element que queda per sota de la diagonal.

Eliminació Gaussiana i Descomposició $LU(VI)$

Lema

Les matrius de la forma $L_1 = \begin{pmatrix} 1 & 0 \\ a_{21} & 1 \end{pmatrix}$, $L_2 = \begin{pmatrix} 1 & 0 \\ -a_{21} & 1 \end{pmatrix}$

són inverses l'una de l'altra, e.d., $L_1 L_2 = L_2 L_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = Id_2.$

Així, si anomenem U a la matriu triangular superior que resulta del procés,

$$U = \begin{pmatrix} 3 & 1 \\ 0 & -4/3 \end{pmatrix}$$

també podem escriure

$$M_1 A = U \quad \Rightarrow \quad A = M_1^{-1} U = LU = \begin{pmatrix} 1 & 0 \\ 1/3 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & -4/3 \end{pmatrix}$$

El procés d'eliminació gaussiana s'associa, des del punt de vista computacional, al procés de factorització de la matriu del sistema com el producte d'una matriu triangular superior per una matriu triangular inferior, $A = LU$. La normalització escollida fa que la matriu L tinga uns en la diagonal principal, i es calcule simplement a partir dels **multiplicadors** utilitzats en el procés d'eliminació recursiva.

Eliminació Gaussiana i Descomposició $LU(VII)$

Exemple 2 Considerem el següent sistema 3×3, i la seua corresponent notació matricial,

$$\left\{ \begin{array}{rcl} 3x + y + z & = & 1, \\ x - y - z & = & 0 \\ x + y - z & = & 2 \end{array} \right\}, \quad A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Realitzem el procés d'eliminació gaussiana *normalitzat*:

Pas 1: Eliminar x de totes les equacions menys la primera $\Rightarrow$

- **(nova Eq. 2) = (Eq. 2) + (-1/3) x (Eq. 1)**
- **(nova Eq. 3) = (Eq. 3) + (-1/3) x (Eq. 1)**

Aquests dos passos són equivalents a la pre-multiplicació per la matriu:

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix}$$

per tant: $AX = b \quad \equiv \quad M_1AX = M_1b \quad \equiv \quad A^{(1)}X = b^{(1)}$ amb

$$A^{(1)} = M_1A = \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 \\ 0 & -4/3 & -4/3 \\ 0 & 2/3 & -4/3 \end{pmatrix},$$

$$b^{(1)} = M_1b = \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1/3 \\ 5/3 \end{pmatrix}$$

- El nou sistema $A^{(1)}X = b^{(1)}$ és equivalent a l'anterior, ja que les noves equacions són una **combinació lineal** de les equacions del sistema original.

$$\begin{pmatrix} 3 & 1 & 1 \\ 0 & -4/3 & -4/3 \\ 0 & 2/3 & -4/3 \end{pmatrix} X = \begin{pmatrix} 1 \\ -1/3 \\ -2/3 \end{pmatrix}, \quad \equiv \quad \left\{ \begin{array}{rcrcrcr} 3x & + & y & + & z & = & 1 \\ & - & (4/3)y & - & (4/3)z & = & -1/3 \\ & & (2/3)y & - & (1/3)z & = & -2/3 \end{array} \right\}$$

Eliminació Gaussiana i Descomposició $LU(VIII)$

Des del punt de vista matricial, l'eliminació de la variable x de la segona i la tercera equacions, equival a convertir en zero els elements de la primera columna per sota de la diagonal principal. Això s'aconsegueix pre-multiplicant per la matriu M_1 que és una matriu triangular inferior, amb uns en la diagonal i amb els **multiplicadors** apropiats per a eliminar la variable de la segona i tercera equació, davall de la diagonal principal de la primera columna.

En el segon pas del procés, es busca eliminar la variable y de la tercera equació o, en termes matricials, es busca aconseguir un zero per sota de la diagonal de la segona columna. Seguint el criteri de normalització anterior, hem de

- multiplicar la segona fila per $-(2/3)/(-4/3) = 1/2$ i sumar a la tercera, deixant el resultat en la tercera. Això s'aconsegueix amb la matriu

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}$$

$$M_2 A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 \\ 0 & -4/3 & -4/3 \\ 0 & 2/3 & -4/3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 \\ 0 & -4/3 & -4/3 \\ 0 & 0 & -2 \end{pmatrix} = A^{(2)}$$

$$b^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1/3 \\ -2/3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1/3 \\ -5/6 \end{pmatrix}.$$

De nou, el sistema $A^{(2)}X = b^{(2)}$

$$\begin{pmatrix} 3 & 1 & 1 \\ 0 & -4/3 & -4/3 \\ 0 & 0 & -2 \end{pmatrix} X = \begin{pmatrix} 1 \\ -1/3 \\ -5/6 \end{pmatrix}, \qquad \equiv \qquad \left\{ \begin{array}{rcrcrcr} 3x & + & y & + & z & = & 1 \\ & - & (4/3)y & - & (4/3)z & = & -1/3 \\ & & & - & 2z & = & -5/6 \end{array} \right\}$$

és equivalent a $A^{(1)}X = b^{(1)}$, i per tant al sistema original $AX = b$.

Eliminació Gaussiana i Descomposició $LU(IX)$

El procés d'eliminació gaussiana per aquest sistema es pot descriure matricialment com a

$$AX = b \quad \equiv \quad \underbrace{M_1 A}_{A^{(1)}} X = \underbrace{M_1 b}_{b^{(1)}} \quad \equiv \quad \underbrace{M_2 M_1 A}_{A^{(2)}=U} X = \underbrace{M_2 M_1 b}_{b^{(2)}}$$

Observeu que les matrius M_1, M_2 són triangulars inferiors, amb uns en la diagonal i amb una estructura tal que es compleix:

Lema

Si les matrius M_1 i M_2 són matrius triangulars inferiors, amb uns en la diagonal i de la forma particular:

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & a_{32} & 1 \end{pmatrix}$$

Podeu comprovar que:

1. $M_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -a_{21} & 1 & 0 \\ -a_{31} & 0 & 1 \end{pmatrix}$
2. $M_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -a_{32} & 1 \end{pmatrix}$
3. $M_1 M_2 = \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & a_{32} & 1 \end{pmatrix}$

Eliminació Gaussiana i Descomposició $LU(X)$

La factorització $A = LU$ es dedueix de les propietats d'aquesta classe de matrius. En el nostre cas, podem escriure

$$M_2 M_1 A = U \quad \rightarrow \quad A = (M_2 M_1)^{-1} U = M_1^{-1} M_2^{-1} U = LU$$

En l'exemple 2 tenim

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \quad \rightarrow \quad M_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{pmatrix}$$

i

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{pmatrix} \quad \rightarrow \quad M_2^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}$$

A més

$$L = M_1^{-1} M_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 1/3 & 1/2 & 1 \end{pmatrix}$$

Els elements de L estan directament relacionats amb els **multiplicadors** utilitzats en l'eliminació gaussiana.

Exercici 1: *Calculeu la descomposició LU de la següent matriu:*

$$A = \begin{pmatrix} 2 & 3 & -1 \\ 4 & 5 & -3 \\ -2 & 3 & -1 \end{pmatrix}$$

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

La descomposició LU i la solució de sistemes lineals: bases

¶¶ Siga el sistema lineal $AX = b$ y $A = LU$, podem escriure

$$AX = b \qquad \equiv \qquad (LU)X = b \qquad \equiv \qquad L(UX) = b$$

La solució del sistema, X, s'aconsegueix resolent dos sistemes lineals amb matriu triangular:

1) $LZ = b$,
2) $UX = Z$

Lògicament, el cost computacional és $\mathcal{N} \sim 2n^2$ (enfront de $\mathcal{N} \sim n^3$ de l'eliminació gaussiana!!).

¶ **Cost computacional:** El nombre d'operacions: $\mathcal{N} \approx 2n^2$ per a un PC típic de 10 gigaflops:

- $n = 16 \implies T_{\text{calcul}} \approx 2 \cdot 16^2/10^{10} = 512/10^{10} \approx 5 \cdot 10^{-8}$ s
- $n = 20 \implies T_{\text{calcul}} \approx 8 \cdot 10^{-8}$ s
- $n = 100 \implies T_{\text{calcul}} \approx 2 \cdot 10^{-6}$ s

La descomposició LU i la solució de sistemes lineals: exemple

Mostrarem el procés de resolució amb un exemple de 3 equacions amb 3 incògnites:

$$\left\{\begin{array}{rcl} 3x+5y+2z &=& 3, \\ x+5y-z &=& 0 \\ x+y-z &=& 2 \end{array}\right\}, \quad A=\begin{pmatrix}3&5&2\\1&5&-1\\1&1&-1\end{pmatrix} \quad X=\begin{pmatrix}x\\y\\z\end{pmatrix} \quad b=\begin{pmatrix}3\\0\\2\end{pmatrix}.$$

¶ Primer fem la descomposició LU de A:

$$\begin{pmatrix}3&5&2\\1&5&-1\\1&1&-1\end{pmatrix} \Rightarrow \begin{pmatrix}\text{mantenim la 1a fila}\\ \text{1fil x } (-1/3) + \text{2fil}\\ \text{1fil x } (-1/3) + \text{3fil}\end{pmatrix} \Rightarrow \begin{pmatrix}3&5&2\\0&10/3&-5/3\\0&-2/3&-5/3\end{pmatrix} \equiv A^{(1)}$$

$$\begin{pmatrix}3&5&2\\0&10/3&-5/3\\0&-2/3&-5/3\end{pmatrix} \Rightarrow \begin{pmatrix}\text{mantenim la 1a fila}\\ \text{mantenim la 2a fila}\\ \text{2fil x } (1/5) + \text{3fil}\end{pmatrix} \Rightarrow \begin{pmatrix}3&5&2\\0&10/3&-5/3\\0&0&-2\end{pmatrix} \equiv U$$

$$L=\begin{pmatrix}1&0&0\\1/3&1&0\\1/3&-1/5&1\end{pmatrix}$$

La descomposició LU i la solució de sistemes lineals: exemple

Mostrarem el procés de resolució amb un exemple de 3 equacions amb 3 incògnites:

$$\left\{ \begin{array}{rcl} 3x+5y+2z & = & 3, \\ x+5y-z & = & 0 \\ x+y-z & = & 2 \end{array} \right\}, \quad A = \begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ Ara resolem el primer dels 2 sistemes (molt més senzill que l'original):

$$LZ = b \Rightarrow \left\{ \begin{array}{rcl} z_1 & = & 3 \\ z_1/3 + z_2 & = & 0 \\ z_1/3 - z_2/5 + z_3 & = & 2 \end{array} \right\} \Rightarrow Z = \begin{pmatrix} 3 \\ -1 \\ 4/5 \end{pmatrix}.$$

¶ Finalment, resolem el segon dels 2 sistemes (molt més senzill que l'original):

$$UX = Z \Rightarrow \left\{ \begin{array}{rcl} 3x+5y+2z & = & 3 \\ \frac{10}{3}y - \frac{5}{3}z & = & -1 \\ -2z & = & 4/5 \end{array} \right\} \Rightarrow X = \begin{pmatrix} 21/10 \\ -1/2 \\ -2/5 \end{pmatrix}.$$

Podeu comprovar que la solució trobada es correspon amb la del sistema original $AX = b$.

La descomposició LU i la solució de sistemes lineals (VI)

Exemple 1: Resoldre el sistema

$$\left\{ \begin{array}{rcl} 3x+y & = & 2 \\ x-y & = & 0 \end{array} \right\};$$

sabent que la matriu de coeficients

$$A = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$$

es pot escriure com $A = LU$ amb **(Noteu que la descomposició NO és única!)**

$$a) \quad L = \begin{pmatrix} 1 & 0 \\ 1/3 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & 1 \\ 0 & -4/3 \end{pmatrix} \qquad b) \quad L = \begin{pmatrix} 1 & 0 \\ 1/3 & -1/3 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix}$$

Exemple 2 Resoldre el sistema

$$\left\{ \begin{array}{rcl} 3x+y+z & = & 1, \\ x-y-z & = & 0 \\ x+y-z & = & 2 \end{array} \right\},$$

sabent que la matriu de coeficients

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$$

es pot escriure com $A = LU$ amb

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 1/3 & -1/2 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 3 & 1 & 1 \\ 0 & -4/3 & -4/3 \\ 0 & 0 & -2 \end{pmatrix}$$

¶¶ *Si considerem el sistema $AX = c$, amb la mateixa matriu de coeficients, però diferent costat dret, no cal tornar a realitzar el procés d'eliminació gaussiana amb el nou sistema. La solució d'aquest nou sistema es calcula de la mateixa manera, resolent els sistemes triangulars*

$$LZ = c, \qquad UX = Z$$

Exercici 2: Calculeu la solució dels sistemes

$$AX = e_1; \qquad AX = e_2; \qquad AX = e_3$$

on $e_1 = (1,0,0)^T$, $e_2 = (0,1,0)^T$; $e_3 = (0,0,1)^T$.

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

El càlcul de la inversa d'una matriu. Càlcul de determinants.

¶¶ La manera més eficient de calcular **la inversa d'una matriu** és resolent sistemes lineals. Si denotem per AI_i la columna i-èsima de la matriu A^{-1}, es compleix que

$$AA^{-1} = Id_n \quad \rightarrow \quad AAI_i = e_i, \quad i = 1, \dots, n$$

on e_i és la i-èsima columna de la matriu identitat, Id_n, és a dir, les components són totes nul·les excepte la i-èsima. Tenint en compte la relació anterior, podem calcular la descomposició LU de A i utilitzar-la per a resoldre els n sistemes lineals

$$AX = e_i, \quad i = 1, \dots, n$$

les solucions d'aquests sistemes ens donen les columnes de la matriu A^{-1}.

¶¶ **El determinant d'una matriu** també es pot deduir de la descomposició LU d'una manera molt eficient. En efecte, si

$$A = LU \quad \rightarrow \quad |A| = |L||U|$$

d'altra banda, per a matrius triangulars es té el següent resultat:

¶¶ *El determinant d'una matriu triangular és el producte dels elements de la diagonal principal.*

Així, si L és una matriu triangular inferior amb uns en la diagonal, tindrem que

$$|L| = 1; \qquad |U| = u_{11} \cdot u_{22} \cdots u_{nn} \quad \rightarrow \quad |A| = u_{11} \cdot u_{22} \cdots u_{nn}$$

Exercici 3: *Usant la descomposició LU, calcula el determinant i la inversa de la matriu*

$$A = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 1 & -3 \\ 2 & 4 & 0 \end{pmatrix}$$

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

Pivotatge i matrius permutació: descripció del problema

¶¶ Com s'ha descrit, *el procés d'eliminació gaussiana no es pot completar si en el pas k, la variable x_k no apareix en la k-èsima equació del sistema*.

Exemple:

$$\begin{array}{ccccccc} & & x_2 & + & x_3 & = & 1 \\ x_1 & & & + & x_3 & = & 1 \\ x_1 & + & x_2 & & & = & 1 \end{array}$$

La primera equació té un zero en la diagonal, i per tant no es pot utilitzar com a equació pivotal en el procés d'eliminació gaussiana. Com que el coeficient de x_1 en la segona equació no és zero, podem intercanviar les dues primeres equacions. Utilitzant notació matricial, volem passar d'un sistema amb matriu A, a un altre equivalent, amb matriu $\tilde{A}$:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \rightarrow \quad \tilde{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

El pas entre les dues matrius es pot fer pre-multiplicant per la **matriu permutació** P_1

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P_1 A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \tilde{A}$$

Observeu que les files de la matriu P_1 són els vectors e_2, e_1, e_3, de manera que l'efecte de pre-multiplicar per aquesta matriu és el d'intercanviar les files primera i segona.

Pivotatge i matrius permutació

¶¶ En general, el procés d'eliminació gaussiana ha d'incorporar una *estratègia de pivotatge*, quan la incògnita a eliminar no apareix a l'equació apropiada. Des del punt de vista matricial, **l'eliminació de Gauss amb pivotatge** produeix tres matrius: 1 **matriu permutació** P (representa l'estratègia de pivotatge), 1 **matriu triangular inferior** amb uns en la diagonal (conté els multiplicadors utilitzats en l'eliminació), i una **matriu triangular superior** (coeficients de la matriu del sistema triangular superior equivalent a l'inicial).

La descomposició matricial resultant es pot escriure com: $PA = LU$

En l'exemple anterior

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \rightarrow \quad P_1 A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \rightarrow \quad M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$A^{(1)} = M_1 P_1 A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = A^{(2)}$$

En el segon pas no és necessari efectuar cap permutació, així que simplement calculem M_2

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

$$A^{(3)} = M_2 A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = U$$

Pivotatge i matrius permutació

Resumint, i raonant com en els exemples anteriors, tenim

$$M_2 M_1 P_1 A = U \quad \rightarrow \quad P_1 A = (M_2 M_1)^{-1} U = M_1^{-1} M_2^{-1} U = LU$$

És a dir, aquest procés d'eliminació gaussiana ens ha portat a una descomposició matricial de tipus

$$PA = LU$$

amb

$$PA = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad U = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{pmatrix}$$

Comproveu que la igualtat $PA = LU$ és certa (només heu de multiplicar les matrius!).

Pivotatge i matrius permutació

¶¶ Considerem ara el **cas general**, en què es busca la descomposició LU d'una matriu $A \in \mathbb{R}^{n\times n}$, admetent que és possible que hàgem de fer un canvi de files (pivotatge) en passos intermedis. En aquest cas, seguint la notació dels exemples anteriors escrivim:

- $A^{(k-1)}$ matriu després de realitzar $k-1$ passos de l'eliminació gaussiana, és a dir després de fer zeros sota les diagonals de les files $1, 2, \dots, k-1$.
- P_k matriu permutació que canvia una de les files $i = k+1, k+2, \dots, n$ amb la fila k, amb l'objectiu que el coeficient de la incògnita x_k siga diferent de zero. Si no cal canviar files en aquest moment, podem escriure $P_k = Id_n$.
- M_k matriu triangular inferior amb uns en la diagonal que aconsegueix eliminar la incògnita x_k de les equacions $k+1$, $k+2, \dots, n$.

Amb aquesta notació, el pas k de l'eliminació gaussiana es pot escriure com a:

$$A^{(k)} = M_k P_k A^{(k-1)}$$

per tant, el procés complet es pot escriure com a:

$$U = A^{(n)} = M_{n-1}P_{n-1}M_{n-2}P_{n-2}\cdots M_2P_2M_1P_1A$$

Utilitzem que:

1. $M_{n-1}P_{n-1}M_{n-2}P_{n-2}\cdots M_2P_2M_1P_1 = (M_{n-1}M_{n-2}\cdots M_2M_1)(P_{n-1}P_{n-2}\cdots P_2P_1)$
2. $P_{n-1}P_{n-2}\cdots P_2P_1 = P \Longleftarrow$ és una matriu de permutació.

Pivotatge i matrius permutació

I arribem a

$$\begin{aligned} U = A^{(n)} &= M_{n-1}P_{n-1}M_{n-2}P_{n-2}\cdots M_2P_2M_1P_1A \\ &= (M_{n-1}M_{n-2}\cdots M_2M_1)(P_{n-1}P_{n-2}\cdots P_2P_1)A \\ &= (M_{n-1}M_{n-2}\cdots M_2M_1)PA \end{aligned} \tag{26}$$

o equivalentment

$$(M_{n-1}M_{n-2}\cdots M_2M_1)^{-1}U = PA \quad \rightarrow \quad PA = M_1^{-1}M_2^{-1}\ldots M_{n-1}^{-1}U = LU$$

on hem utilitzat el següent resultat, que generalitza el vist en un dels exercicis anteriors respecte a l'estructura del producte de les matrius M_k, per al qual tenim:

El producte $M_1^{-1}\ldots M_{n-1}^{-1}$ és una matriu triangular inferior amb uns en la diagonal.

Pivotatge i matrius permutació

IMPORTANT

No només és necessari utilitzar una estratègia de permutació quan el pivot és nul. També per raons de precisió numèrica quan el pivot és molt xicotet.

Exemple:

$$\begin{aligned} 0.003x_1 + 59.14x_2 &= 59.17 \\ 5.291x_1 - 6.130x_2 &= 46.78 \end{aligned} \; ; \quad \text{solució } \textit{correcta:} \; \begin{aligned} x_1 &= 10 \\ x_2 &= 1 \end{aligned}$$

amb Gauss sense pivotatge tenim:

$$m_{21} = \frac{5.291}{0.00300} = 1763.\bar{6}$$

i ara fent el pas d'eliminació i

$$\begin{aligned} 0.003x_1 + \qquad\qquad 59.14x_2 &= 59.17 \\ -104297.116666864x_2 &= 104356.156666686 \end{aligned} \; ; \quad \text{solució } \begin{aligned} x_1 &= -1.1592 \\ x_2 &\simeq 1.00057 \end{aligned}$$

[vegeu l'enorme error amb què es calcula x_1]

Per a pensar: Com podríem fer servir el procés de pivotatge per a millorar la precisió de l'anterior resultat?

Continguts

1. Resolució algebraica
2. Mètodes directes: Gauss
3. Eliminació gaussiana i descomposició **LU**
4. La descomposició **LU** i la solució de sistemes lineals
5. El càlcul de la inversa d'una matriu. Càlcul de determinants
6. Pivotatge i matrius permutació
7. Mètodes iteratius.

Mètodes Iteratius (I): Introducció

¶¶ En problemes *xicotets* sol ser suficient utilitzar mètodes directes. Però per a sistemes lineals *grans*, especialment aquells que involucren *matrius disperses (sparse)*, hem de saber que *el mètode de Gauss presenta les següents dificultats:*

- És **car**: $\mathcal{O}(n^3)$.
- És **destructiu**: retoca la matriu del sistema, i això pot tenir conseqüències molt poc desitjables (sobretot en matrius disperses). A més requereix en moltes ocasions guardar una còpia de la matriu.
- Per a obtenir una solució cal realitzar tot el procés: en passos intermedis, **no es disposa de cap aproximació de la solució**.

Mètodes Iteratius (I): Introducció

¶¶ En problemes *xicotets* sol ser suficient utilitzar mètodes directes. Però per a sistemes lineals *grans*, especialment aquells que involucren *matrius disperses (sparse)*, hem de saber que *el mètode de Gauss presenta les següents dificultats:*

- És **car**: $\mathcal{O}(n^3)$.
- És **destructiu**: retoca la matriu del sistema, i això pot tenir conseqüències molt poc desitjables (sobretot en matrius disperses). A més requereix en moltes ocasions guardar una còpia de la matriu.
- Per a obtenir una solució cal realitzar tot el procés: en passos intermedis, **no es disposa de cap aproximació de la solució**.

¶¶ Els mètodes iteratius s'introdueixen com un intent de salvar aquestes dificultats.

Un mètode iteratiu pren un sistema lineal $A\mathbf{x} = \mathbf{b}$ i construeix una successió de vectors de manera que

$$\mathbf{x}_m \longrightarrow \mathbf{x}, \text{ quan } m \longrightarrow \infty.$$

La forma de construir aquesta successió depèn de A, $\mathbf{b}$ i del vector inicial.

Mètodes Iteratius (II): Avantatges i inconvenients

¶¶ **Inconvenients:**

- En general, **la convergència pot ser lenta o, simplement, no convergeixen**.
- **Els resultats teòrics sobre convergència són sovint pobres**: hi ha convergència en situacions molt més generals del que la teoria prediu. A més, revisar aquestes hipòtesis pot resultar tan costós com resoldre el mateix sistema.

Mètodes Iteratius (II): Avantatges i inconvenients

¶¶ **Inconvenients:**

- En general, **la convergència pot ser lenta o, simplement, no convergeixen**.
- **Els resultats teòrics sobre convergència són sovint pobres**: hi ha convergència en situacions molt més generals del que la teoria prediu. A més, revisar aquestes hipòtesis pot resultar tan costós com resoldre el mateix sistema.

¶¶ **Avantatges:**

- Mètodes **no destructius** $\Longrightarrow$ No modifiquen la matriu del sistema i, en general, requereixen només multiplicar per la matriu del sistema o per parts d'aquesta.
- Solen ser **més estables** davant els errors d'arrodoniment.
- **Es disposa en cada pas d'una aproximació de la solució**.

Mètodes Iteratius (III): Coneixements previs

¶¶ **Normes vectorials:** Serveixen per a mesurar la *mida* de vectors i matrius. Siga el vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$. Es poden definir les següents normes:

Definició analítica

$$\begin{aligned} ||\mathbf{x}||_1 &:= |x_1| + |x_2| + \ldots + |x_n|, \\ ||\mathbf{x}||_2 &:= \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}, \\ ||\mathbf{x}||_\infty &:= \max_{i=1,\ldots,n} |x_i|. \end{aligned}$$

Ordre Matlab

```
norm(a,1)               % norma 1
norm(a,2); norm(a)      % norma 2
norm(a,inf)             % norma infinit
```

¶¶ Cada norma vectorial defineix alhora una **norma matricial induïda**

$$||A|| = \sup_{||\mathbf{x}||=1} ||A\mathbf{x}|| \quad \Longrightarrow \quad \begin{aligned} ||A||_1 &= \max_{j=1,\ldots,n} \sum_{i=1}^{n} |a_{ij}|, \\ ||A||_\infty &= \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{ij}|, \\ ||A||_2 &= \sqrt{\rho(A^T A)}, \end{aligned}$$

essent $\rho(B)$ el **radi espectral** de la matriu B, és a dir, el més gran dels valors absoluts dels valors propis de B. (Si A és simètrica, $||A||_2 = \rho(A)$).

Mètodes Iteratius (IV): Definició i condicions de convergència (I)

¶¶ Els mètodes iteratius lineals comencen considerant una partició de A tal que

$$A = M - N$$

amb M invertible. Així, si $\mathbf{x}$ és solució de $A\mathbf{x} = \mathbf{b}$,

$$M\mathbf{x} = N\mathbf{x} + \mathbf{b}.$$

El mètode iteratiu consisteix a:

- Prendre com $\mathbf{x}_0$ una aproximació de la solució (c.a., $\mathbf{x}_0 = \mathbf{0}$ si és desconeguda).
- Resoldre en cada pas m

$$M\mathbf{x}_{m+1} = N\mathbf{x}_m + \mathbf{b}. \tag{27}$$

És fàcil veure que si la successió construïda (27) convergeix a algun $\mathbf{x}$, llavors el vector és la solució del sistema. Encara més, es comprova fàcilment que

$$\mathbf{x}_{m+k+1} - \mathbf{x} = (M^{-1}N)(\mathbf{x}_{m+k} - \mathbf{x}) = \ldots = (M^{-1}N)^k(\mathbf{x}_{m+1} - \mathbf{x}).$$

Denotant per $B = M^{-1}N$ es té el següent resultat

Teorema

El mètode iteratiu convergeix sii $\rho(B) < 1$.

Mètodes Iteratius (IV): Definició i condicions de convergència (II)

¶¶ Noteu que la convergència del mètode passa siga quin siga $\mathbf{x}_0$, és a dir, independentment de com és arrancat el mètode i de quin siga el terme independent. Clarament si es té una estimació bona de la solució, el mètode convergirà en menys iteracions, però no és una condició imprescindible per a assegurar la convergència.

¶¶ **Elecció de M**

- Si $M = A$ aleshores $N = 0 \Longrightarrow B = 0 \Longrightarrow \rho(B) = 0$. És a dir, hi ha convergència en una única iteració. Aquesta iteració consisteix a resoldre directament el sistema d'equacions.
- En general, un assegura la convergència quan M recull la informació més important d'A, de manera que $N = M - A$ tinga una mida petita comparada amb M.
- S'ha de tenir en compte la definició del mètode (27) i que per tant en cada iteració cal resoldre un sistema d'equacions. Així, interessa que M siga senzilla des del punt de vista de la resolució del sistema lineal (c.a., diagonal, triangular...) i que simultàniament reculla la màxima informació possible de A.

Mètodes Iteratius (IV): Resum de condicions de convergència

Verificació de les condicions de convergència d'un mètode iteratiu

- Es descompon $A = M - N$.
- Es calcula $B = M^{-1}N$.
- Convergència sii $\rho(B) < 1$
  ($\rho(B)$ és el més gran dels autovalors de B en valor absolut).

Mètode de Jacobi (I)

¶¶ Mètode de Jacobi:
Donat el sistema $A\mathbf{x} = \mathbf{b}$, amb

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Consisteix a descompondre la matriu A en la seua part diagonal D, una matriu triangular superior U i una triangular inferior L,

$$A = D + L + U$$

$$D = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}, L = \begin{pmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{pmatrix}, U = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Mètode de Jacobi (II)

¶¶ Aplicant l'anterior descomposició al sistema $A\mathbf{x} = \mathbf{b}$ dóna

$$D\mathbf{x} + (L+U)\mathbf{x} = \mathbf{b} \Rightarrow \mathbf{x} = D^{-1}[\mathbf{b} - (L+U)\mathbf{x}]$$

sempre que no tinguem cap element nul en la diagonal, quedant el mètode iteratiu com a:

$$\mathbf{x}_{k+1} = D^{-1}[\mathbf{b} - (L+U)\mathbf{x}_k]$$

Mètode de Jacobi

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left[b_i - \sum_{\substack{j\neq i}}^{n} a_{ij}x_j^{(k)}\right], \quad i = 1, \ldots, n \tag{28}$$

¶ Exemple: Aplicar el mètode de Jacobi al sistema (11).

¶ **Utilització i convergència:** El mètode es pot utilitzar quan la matriu no tinga zeros a la diagonal i convergeix sempre que la matriu siga *diagonal dominant*. Fins i tot pot convergir, en cas de no ser diagonal dominant, si els elements de la diagonal són més grans (en magnitud) que els altres elements.

Definició:

Una matriu és diagonal dominant per files si: $|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|, \quad \forall i$

per columnes si: $|a_{jj}| \geq \sum_{i\neq j} |a_{ij}|, \quad \forall j.$

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzàrem per a sistemes:

$$\left\{ \begin{array}{rcl} 3x_1 + 5x_2 + 2x_3 & = & 3, \\ x_1 + 5x_2 - x_3 & = & 0 \\ x_1 + x_2 - x_3 & = & 2 \end{array} \right\}, \quad A = \begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ Ara identifiquem les matrius D, L i U associades a la matriu A:

$$A = D+L+U = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 5 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}; D^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

¶ Apliquem la fórmula $\mathbf{x}_{k+1} = D^{-1}[\mathbf{b} - (L+U)\mathbf{x}_k]$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & -1 \end{pmatrix} \left[\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 & 5 & 2 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_0 \right]$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 = \underbrace{\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}}_{\text{fix}} - \underbrace{\begin{pmatrix} 0 & 5/3 & 2/3 \\ 1/5 & 0 & -1/5 \\ -1 & -1 & 0 \end{pmatrix}}_{D^{-1}(L+U):\ \text{fix}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_0$$

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzarem per a sistemes:

$$\left\{ \begin{array}{rcl} 3x_1 + 5x_2 + 2x_3 & = & 3, \\ x_1 + 5x_2 - x_3 & = & 0 \\ x_1 + x_2 - x_3 & = & 2 \end{array} \right\}, \quad A = \begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ La iteració comença donant valors a $\mathbf{x}_0$. Podem començar fent $\mathbf{x}_0 = [0, 0, 0]^T$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 1/5 & 0 & -1/5 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

¶ Avaluem l'error que hem comès i comprovem si és menor que una certa tolerància. Una possibilitat (encara que no l'única) és avaluar la norma de la diferència entre dues iteracions consecutives:

$$\text{err}_1 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_0 \right\| = \left\| \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \right\| = \sqrt{1^2 + 0^2(-2)^2} = \sqrt{5}$$

¶ Com que $\text{err}_1 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzarem per a sistemes:

$$\left\{ \begin{array}{rcl} 3x_1 + 5x_2 + 2x_3 & = & 3, \\ x_1 + 5x_2 - x_3 & = & 0 \\ x_1 + x_2 - x_3 & = & 2 \end{array} \right\}, \quad A = \begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ La segona iteració la fem utilitzant el valor calculat de $\mathbf{x}_1$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 1/5 & 0 & -1/5 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} -4/3 \\ 3/5 \\ -1 \end{pmatrix} = \begin{pmatrix} 7/3 \\ -3/5 \\ -1 \end{pmatrix}$$

¶ Tornem a avaluar l'error comès

$$\text{err}_2 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_2 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 \right\| = \left\| \begin{pmatrix} 7/3 \\ -3/5 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 4/3 \\ -3/5 \\ 1 \end{pmatrix} \right\| = 1.771$$

¶ Com que $\text{err}_2 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzarem per a sistemes:

$$\left\{\begin{array}{rcl} 3x_1+5x_2+2x_3 & = & 3, \\ x_1+5x_2-x_3 & = & 0 \\ x_1+x_2-x_3 & = & 2 \end{array}\right\}, \quad A=\begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b=\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ La tercera iteració la fem utilizant el valor calculat de $\mathbf{x}_2$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_3 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 1/5 & 0 & -1/5 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 7/3 \\ -3/5 \\ -1 \end{pmatrix}_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} -5/3 \\ 2/3 \\ -26/15 \end{pmatrix} = \begin{pmatrix} 8/3 \\ -2/3 \\ -4/15 \end{pmatrix}$$

¶ Tornem a avaluar l'error comès

$$\mathrm{err}_3 = \left\|\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_3 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_2\right\| = \left\|\begin{pmatrix} 8/3 \\ -2/3 \\ -4/15 \end{pmatrix} - \begin{pmatrix} 7/3 \\ -3/5 \\ -1 \end{pmatrix}\right\| = \left\|\begin{pmatrix} 1/3 \\ -1/15 \\ 11/15 \end{pmatrix}\right\| = 0.808$$

¶ Com que $\mathrm{err}_3 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzarem per a sistemes:

$$\left\{\begin{array}{rcl} 3x_1+5x_2+2x_3 & = & 3, \\ x_1+5x_2-x_3 & = & 0 \\ x_1+x_2-x_3 & = & 2 \end{array}\right\}, \quad A=\begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b=\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ La quarta iteració la fem utilizant el valor calculat de $\mathbf{x}_3$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_4 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 1/5 & 0 & -1/5 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 8/3 \\ -2/3 \\ -4/15 \end{pmatrix}_3 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} -58/45 \\ -44/75 \\ -2 \end{pmatrix} = \begin{pmatrix} 103/45 \\ -44/75 \\ 0 \end{pmatrix}$$

¶ Tornem a avaluar l'error comès

$$\mathrm{err}_4 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_4 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_3 \right\| = \left\| \begin{pmatrix} 103/45 \\ -44/75 \\ 0 \end{pmatrix} - \begin{pmatrix} 8/3 \\ -2/3 \\ -4/15 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 17/45 \\ 2/25 \\ 4/15 \end{pmatrix} \right\| = 0.4693$$

¶ Com que $\mathrm{err}_4 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzarem per a sistemes:

$$\left\{\begin{array}{rcl} 3x_1+5x_2+2x_3 & = & 3, \\ x_1+5x_2-x_3 & = & 0 \\ x_1+x_2-x_3 & = & 2 \end{array}\right\}, \quad A=\begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b=\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ La quinta iteració la fem utilizant el valor calculat de $\mathbf{x}_4$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_5 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 1/5 & 0 & -1/5 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 103/45 \\ -44/75 \\ 0 \end{pmatrix}_4 = \begin{pmatrix} 89/45 \\ -103/225 \\ -67/225 \end{pmatrix}$$

¶ Tornem a avaluar l'error comès

$$\text{err}_5 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_5 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_4 \right\| = \left\| \begin{pmatrix} 89/45 \\ -103/225 \\ -67/225 \end{pmatrix} - \begin{pmatrix} 103/45 \\ -44/75 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -14/45 \\ 29/225 \\ -67/225 \end{pmatrix} \right\| = 0.4495$$

¶ Com que $\text{err}_5 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Jacobi (III): Exemple.

Mostrarem el procés de resolució iteratiu de Jacobi per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que utilitzarem per a sistemes:

$$\left\{\begin{array}{rcl} 3x_1+5x_2+2x_3 &=& 3, \\ x_1+5x_2-x_3 &=& 0 \\ x_1+x_2-x_3 &=& 2 \end{array}\right\}, \quad A=\begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b=\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ Repetint el procés fins a la iteració 9 obtenim:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_9 = \begin{pmatrix} 2.1306 \\ -0.5105 \\ -0.3715 \end{pmatrix}; \quad \mathsf{err}_9 = 0.0404$$

¶ Com que $\mathsf{err}_9 < 10^{-1}$, concloem el procés iteratiu.

¶ Podem comparar el resultat del procés iteratiu amb la solució exacta del problema (que ja havíem calculat prèviament):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\mathsf{exacta}} = \begin{pmatrix} 2.1 \\ -0.5 \\ -0.4 \end{pmatrix}; \quad \mathsf{error} = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\mathsf{exacta}} - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_9 \right\| = \left\| \begin{pmatrix} -0.0306 \\ 0.0105 \\ -0.0285 \end{pmatrix} \right\| = 0.0431$$

Mètode de Gauss-Seidel

¶¶ Podem millorar l'algorisme de Jacobi observant que en la fórmula (28) per a calcular $x_i^{(k+1)}$, utilitzem les components de la iteració anterior $\mathbf{x}^{(k)}$. Atès que per a $i > 1$, ja s'han calculat els valors $x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_{i-1}^{(k+1)}$, es pot esperar una convergència més ràpida a la solució si fem servir els valors ja calculats en la iteració $k+1$ l'algoritme per a avaluar els valors de $x_i^{(k+1)}, \ldots, x_n^{(k+1)}$. Això ens porta a la següent definició de l'algorisme (suposant com en el cas del mètode de Jacobi que $a_{ii} \neq 0,\ i = 1, \ldots, n$)

Mètode de Gauss-Seidel

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right], \quad i = 1, \ldots, n \tag{29}$$

¶ Exemple: Aplicar el mètode de Gauss-Seidel al sistema (11)

¶ **Utilització i convergència:** Igual que el mètode de Jacobi.

¶ **Nota:** Matricialment, els mètodes de Jacobi i Gauss-Seidel són una mica diferents:

$$\text{Jacobi:} \quad D\mathbf{x}_{k+1} = \mathbf{b} - (L+U)\mathbf{x}_k \tag{30}$$

$$\text{Gauss-Seidel:} \quad (D+L)\mathbf{x}_{k+1} = \mathbf{b} - U\mathbf{x}_k \tag{31}$$

Mètode de Gauss-Seidel (II): Exemple.

Mostrarem el procés de resolució iteratiu de Gauss-Seidel per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que abans:

$$\left\{\begin{array}{rcl} 3x_1+5x_2+2x_3 & = & 3, \\ x_1+5x_2-x_3 & = & 0 \\ x_1+x_2-x_3 & = & 2 \end{array}\right\}, \quad A=\begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b=\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ Ara identifiquem les matrius D, L i U associades a la matriu A:

$$A=D+L+U=\begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{pmatrix}+\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}+\begin{pmatrix} 0 & 5 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}; D^{-1}=\begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

¶ Apliquem la fórmula $\mathbf{x}_{k+1}=(D+L)^{-1}[\mathbf{b}-U\mathbf{x}_k]$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1=\begin{pmatrix} 1/3 & 0 & 0 \\ -1/15 & 1/5 & 0 \\ 4/15 & 1/5 & -1 \end{pmatrix}\left[\begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}-\begin{pmatrix} 0 & 5 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_0\right]$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1=\underbrace{\begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix}}_{\text{fix}}-\underbrace{\begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix}}_{(D+L)^{-1}U:\ \text{fix}}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_0$$

Mètode de Gauss-Seidel (II): Exemple.

Mostrarem el procés de resolució iteratiu de Gauss-Seidel per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que abans:

$$\left\{\begin{array}{rcl} 3x_1 + 5x_2 + 2x_3 & = & 3, \\ x_1 + 5x_2 - x_3 & = & 0 \\ x_1 + x_2 - x_3 & = & 2 \end{array}\right\}, \quad A = \begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ Donem un valor inicial per a començar el procés iteratiu, com ara $\mathbf{x}_0 = [0,0,0]^T$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}_0 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix}$$

¶ Avaluem l'error (igual que en el mètode de Jacobi):

$$\text{err}_1 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_0 \right\| = \left\| \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} \right\| = 1.575$$

¶ Com que $\text{err}_1 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Gauss-Seidel (II): Exemple.

Mostrarem el procés de resolució iteratiu de Gauss-Seidel per a obtenir una solució aproximada amb una tolerància de 10^{-1}, i amb el mateix exemple que abans:

$$\left\{\begin{array}{rcl} 3x_1 + 5x_2 + 2x_3 & = & 3, \\ x_1 + 5x_2 - x_3 & = & 0 \\ x_1 + x_2 - x_3 & = & 2 \end{array}\right\}, \quad A = \begin{pmatrix} 3 & 5 & 2 \\ 1 & 5 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

¶ Fem la segona iteració utilitzant el valor de $\mathbf{x}_1$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_2 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix}_1 = \begin{pmatrix} 2.1333 \\ -0.66667 \\ -0.53333 \end{pmatrix}$$

¶ Avaluem l'error (igual que en el mètode de Jacobi):

$$\text{err}_2 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_2 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_1 \right\| = \left\| \begin{pmatrix} 2.1333 \\ -0.66667 \\ -0.53333 \end{pmatrix} - \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1.1333 \\ -0.46667 \\ 0.66667 \end{pmatrix} \right\| = 1.395$$

¶ Com que $\text{err}_2 > 10^{-1}$, hem de continuar el procés iteratiu.

Mètode de Gauss-Seidel (II): Exemple.

¶ Fem la tercera iteració utilitzant el valor de $\mathbf{x}_2$ i avaluem l'error:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_3 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 2.1333 \\ -0.66667 \\ -0.53333 \end{pmatrix}_2 = \begin{pmatrix} 2.4667 \\ -0.6 \\ -0.13333 \end{pmatrix}$$

$$\text{err}_3 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_3 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_2 \right\| = \left\| \begin{pmatrix} 2.4667 \\ -0.6 \\ -0.13333 \end{pmatrix} - \begin{pmatrix} 2.1333 \\ -0.66667 \\ -0.53333 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0.33333 \\ 0.066667 \\ 0.4 \end{pmatrix} \right\| = 0.525$$

¶ Com que $\text{err}_3 > 10^{-1}$, hem de continuar el procés iteratiu.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_4 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 2.4667 \\ -0.6 \\ -0.13333 \end{pmatrix}_3 = \begin{pmatrix} 2.0889 \\ -0.44444 \\ -0.35556 \end{pmatrix}$$

$$\text{err}_4 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_4 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_3 \right\| = \left\| \begin{pmatrix} 2.0889 \\ -0.44444 \\ -0.35556 \end{pmatrix} - \begin{pmatrix} 2.4667 \\ -0.6 \\ -0.13333 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -0.37778 \\ 0.15556 \\ -0.22222 \end{pmatrix} \right\| = 0.465$$

Mètode de Gauss-Seidel (II): Exemple.

¶ Fem la cinquena iteració utilitzant el valor de $\mathbf{x}_4$ i avaluem l'error:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_5 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 2.0889 \\ -0.44444 \\ -0.35556 \end{pmatrix}_4 = \begin{pmatrix} 1.9778 \\ -0.46667 \\ -0.48889 \end{pmatrix}$$

$$\text{err}_5 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_5 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_4 \right\| = \left\| \begin{pmatrix} 1.9778 \\ -0.46667 \\ -0.48889 \end{pmatrix} - \begin{pmatrix} 2.0889 \\ -0.44444 \\ -0.35556 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -0.11111 \\ -0.02222 \\ -0.13333 \end{pmatrix} \right\| = 0.175$$

¶ Com que $\text{err}_5 > 10^{-1}$, hem de continuar el procés iteratiu.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_6 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1.9778 \\ -0.46667 \\ -0.48889 \end{pmatrix}_5 = \begin{pmatrix} 2.1037 \\ -0.51852 \\ -0.41481 \end{pmatrix}$$

$$\text{err}_6 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_6 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_5 \right\| = \left\| \begin{pmatrix} 2.1037 \\ -0.51852 \\ -0.41481 \end{pmatrix} - \begin{pmatrix} 1.9778 \\ -0.46667 \\ -0.48889 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0.12593 \\ -0.05185 \\ 0.074074 \end{pmatrix} \right\| = 0.155$$

Mètode de Gauss-Seidel (II): Exemple.

¶ Fem la setena iteració utilitzant el valor de $\mathbf{x}_6$ i avaluem l'error:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_7 = \begin{pmatrix} 1 \\ -1/5 \\ -6/5 \end{pmatrix} - \begin{pmatrix} 0 & 5/3 & 2/3 \\ 0 & -1/3 & -1/3 \\ 0 & 4/3 & 1/3 \end{pmatrix} \begin{pmatrix} 2.1037 \\ -0.51852 \\ -0.41481 \end{pmatrix}_6 = \begin{pmatrix} 2.1407 \\ -0.51111 \\ -0.37037 \end{pmatrix}$$

$$\text{err}_7 = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_7 - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_6 \right\| = \left\| \begin{pmatrix} 2.1407 \\ -0.51111 \\ -0.37037 \end{pmatrix} - \begin{pmatrix} 2.1037 \\ -0.51852 \\ -0.41481 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0.037037 \\ 0.007407 \\ 0.044444 \end{pmatrix} \right\| = 0.058$$

¶ Com que $\text{err}_7 < 10^{-1}$, ja podem parar el procés iteratiu.
¶ Podem comparar el resultat del procés iteratiu amb la solució exacta del problema (que ja havíem calculat prèviament):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\text{exacta}} = \begin{pmatrix} 2.1 \\ -0.5 \\ -0.4 \end{pmatrix}; \quad \text{error} = \left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\text{exacta}} - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_7 \right\| = \left\| \begin{pmatrix} -0.0407 \\ 0.0111 \\ -0.0296 \end{pmatrix} \right\| = 0.052$$

¶ En aquest exemple ja es veu que el mètode de Gauss-Seidel convergeix més de pressa que el de Jacobi. Açò és una tendència general.

VNIVERSITAT DE VALÈNCIA

Matemàtiques III: 3. Mètodes numèrics per a la resolució d'equacions no lineals

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1. Introducció

2. Mètode de la bisecció

3. Mètode de Newton

Continguts

1 Introducció

2 Mètode de la bisecció

3 Mètode de Newton

Per a què serveix tot açò?

¶ En la major part de les aplicacions científiques i tècniques és comú haver de trobar el valor d'una variable la dependència funcional de la qual amb una altra és complicada, o fins i tot desconeguda.
Aquesta situació es dóna quan, p. ex.,

- no existeix una funció explícita que ens permeta *invertir* la relació entre dues variables.
  Exemple: Considereu la funció

$$y(x) = 10^6 e^x + \frac{4.35 \cdot 10^5}{x}(e^x - 1),$$

  en aquest cas, no és possible obtenir de manera explícita la relació inversa $x(y)$.

Per a què serveix tot açò?

¶ En la major part de les aplicacions científiques i tècniques és comú haver de trobar el valor d'una variable la dependència funcional de la qual amb una altra és complicada, o fins i tot desconeguda.
Aquesta situació es dóna quan, p. ex.,

- no existeix una funció explícita que ens permeta *invertir* la relació entre dues variables.
  Exemple: Considereu la funció

$$y(x) = 10^6 e^x + \frac{4.35 \cdot 10^5}{x}(e^x - 1),$$

  en aquest cas, no és possible obtenir de manera explícita la relació inversa $x(y)$.

- Una altra possibilitat és que la funció $y(x)$ siga explícita, però els valors de y depenen dels de x de manera molt complicada.
  Exemple: Considereu la funció $y(x) = ax^3 + bx^2 + cx + d$, la qual té solució analítica (però complicada i amb un cost computacional molt alt):

$$x = \frac{\sqrt[3]{\sqrt{\left(-27a^2d + 9abc - 2b^3\right)^2 + 4\left(3ac - b^2\right)^3} - 27a^2d + 9abc - 2b^3}}{3\sqrt[3]{2}a} - \frac{\sqrt[3]{2}\left(3ac - b^2\right)}{3a\sqrt[3]{\sqrt{\left(-27a^2d + 9abc - 2b^3\right)^2 + 4\left(3ac - b^2\right)^3} - 27a^2d + 9abc - 2b^3}} - \frac{b}{3a}$$

Per a què serveix tot açò?

¶ En la major part de les aplicacions científiques i tècniques és comú haver de trobar el valor d'una variable la dependència funcional de la qual amb una altra és complicada, o fins i tot desconeguda.
Aquesta situació es dóna quan, p. ex.,

- no existeix una funció explícita que ens permeta *invertir* la relació entre dues variables.
  Exemple: Considereu la funció

  $$y(x) = 10^6 e^x + \frac{4.35 \cdot 10^5}{x}(e^x - 1),$$

  en aquest cas, no és possible obtenir de manera explícita la relació inversa $x(y)$.

- Una altra possibilitat és que la funció $y(x)$ siga explícita, però els valors de y depenen dels de x de manera molt complicada.
  Exemple: Considereu la funció $y(x) = ax^3 + bx^2 + cx + d$, la qual té solució analítica (però complicada i amb un cost computacional molt alt):

$$x = \frac{\sqrt[3]{\sqrt{\left(-27a^2d + 9abc - 2b^3\right)^2 + 4\left(3ac - b^2\right)^3} - 27a^2d + 9abc - 2b^3}}{3\sqrt[3]{2}a} - \frac{\sqrt[3]{2}\left(3ac - b^2\right)}{3a\sqrt[3]{\sqrt{\left(-27a^2d + 9abc - 2b^3\right)^2 + 4\left(3ac - b^2\right)^3} - 27a^2d + 9abc - 2b^3}} - \frac{b}{3a}$$

En aquests casos podem estar interessats a conèixer el valor de x si el valor de y és conegut. Per a això, podem reformular el problema de manera que es reduïsca a trobar el "zero" d'una funció no lineal. En l'exemple anterior, si tinguérem $y = 1.5 \cdot 10^6$, el problema es reduiria a trobar el zero de la funció

$$0 = f(x) = 10^6 e^x + \frac{4.35 \cdot 10^5}{x}(e^x - 1) - 1.5 \cdot 10^6.$$

Introducció (II)

Bibliografia: Burden & Faires, Cap. 2

Objectiu

Descriure les tècniques numèriques més bàsiques per a trobar les solucions (zeros) de la funció $f(x)$.

Introducció (II)

Bibliografia: Burden & Faires, Cap. 2

Objectiu

Descriure les tècniques numèriques més bàsiques per a trobar les solucions (zeros) de la funció $f(x)$.

¶ Es tracta de resoldre numèricament l'equació:

$$f(x) = 0 \tag{1}$$

Introducció (II)

Bibliografia: Burden & Faires, Cap. 2

Objectiu

Descriure les tècniques numèriques més bàsiques per a trobar les solucions (zeros) de la funció $f(x)$.

¶ Es tracta de resoldre numèricament l'equació:

$$f(x) = 0 \tag{1}$$

¶ Suposarem, en general, que la funció $f(x)$ és **contínua**.

Introducció (II)

Bibliografia: Burden & Faires, Cap. 2

Objectiu

Descriure les tècniques numèriques més bàsiques per a trobar les solucions (zeros) de la funció $f(x)$.

¶ Es tracta de resoldre numèricament l'equació:

$$f(x) = 0 \tag{1}$$

¶ Suposarem, en general, que la funció $f(x)$ és **contínua**.

¶ Aquestes tècniques consisteixen, en general, en **mètodes iteratius** que produeixen una successió de valors aproximats de la solució que s'espera que convergisca a l'**arrel** de l'equació.

Introducció (II)

Bibliografia: Burden & Faires, Cap. 2

Objectiu

Descriure les tècniques numèriques més bàsiques per a trobar les solucions (zeros) de la funció $f(x)$.

¶ Es tracta de resoldre numèricament l'equació:

$$f(x) = 0 \tag{1}$$

¶ Suposarem, en general, que la funció $f(x)$ és **contínua**.

¶ Aquestes tècniques consisteixen, en general, en **mètodes iteratius** que produeixen una successió de valors aproximats de la solució que s'espera que convergisca a l'**arrel** de l'equació.

¶ Criteris per a l'elecció del mètode:

1. la capacitat de separar arrels properes,
2. la seua fiabilitat per a obtenir les solucions dins d'un cert error,
3. l'ordre de convergència (té a veure amb la velocitat amb què arribem a la solució).

Continguts

1 Introducció

2 Mètode de la bisecció

3 Mètode de Newton

Mètode de la bisecció: la base

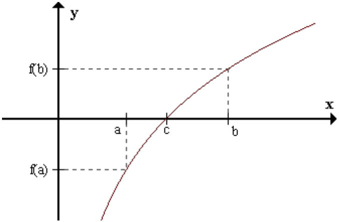
Algorisme de cerca d'arrels que opera dividint l'interval per la meitat i seleccionant el subinterval que té l'arrel. Es basa en l'aplicació del **Teorema de Bolzano**.

Mètode de la bisecció: la base

Algorisme de cerca d'arrels que opera dividint l'interval per la meitat i seleccionant el subinterval que té l'arrel. Es basa en l'aplicació del **Teorema de Bolzano**.

Teorema de Bolzano

Si una funció és contínua en un interval tancat i fitat, i en els seus extrems pren valors amb signes oposats, existeix almenys una arrel de la funció a l'interior de l'interval.



Mètode de la bisecció: Algorisme

¶¶ Siga una funció $f(x)$, contínua en l'interval $[a, b]$, tal que: $f(a) \cdot f(b) < 0$.

- Es calcula el punt mitjà m de l'interval $[a, b]$ i s'avalua $f(m)$. Si aquest valor és igual a zero, ja hem trobat l'arrel buscada. En cas que no ho siga,
- verifiquem si $f(m)$ té signe oposat amb $f(a)$ o amb $f(b)$ i, en conseqüència, es redefineix l'interval $[a, b]$ com $[a, m]$ o $[m, b]$.
- Amb aquest nou interval es continua successivament tancant la solució a un interval cada vegada més petit, fins a aconseguir la precisió desitjada.

Mètode de la bisecció: Algorisme

¶¶ Siga una funció $f(x)$, contínua en l'interval $[a, b]$, tal que: $f(a) \cdot f(b) < 0$.

- Es calcula el punt mitjà m de l'interval $[a, b]$ i s'avalua $f(m)$. Si aquest valor és igual a zero, ja hem trobat l'arrel buscada. En cas que no ho siga,
- verifiquem si $f(m)$ té signe oposat amb $f(a)$ o amb $f(b)$ i, en conseqüència, es redefineix l'interval $[a, b]$ com $[a, m]$ o $[m, b]$.
- Amb aquest nou interval es continua successivament tancant la solució a un interval cada vegada més petit, fins a aconseguir la precisió desitjada.

¶¶ **La bisecció convergeix linealment.**

¶¶ Una cota de l'error absolut és: $\frac{|b-a|}{2^n}$ a la n-èsima iteració.

Mètode de la bisecció: Algorisme

¶¶ Siga una funció $f(x)$, contínua en l'interval $[a,b]$, tal que: $f(a)\cdot f(b)<0$.

- Es calcula el punt mitjà m de l'interval $[a,b]$ i s'avalua $f(m)$. Si aquest valor és igual a zero, ja hem trobat l'arrel buscada. En cas que no ho siga,
- verifiquem si $f(m)$ té signe oposat amb $f(a)$ o amb $f(b)$ i, en conseqüència, es redefineix l'interval $[a,b]$ com $[a,m]$ o $[m,b]$.
- Amb aquest nou interval es continua successivament tancant la solució a un interval cada vegada més petit, fins a aconseguir la precisió desitjada.

¶¶ **La bisecció convergeix linealment.**

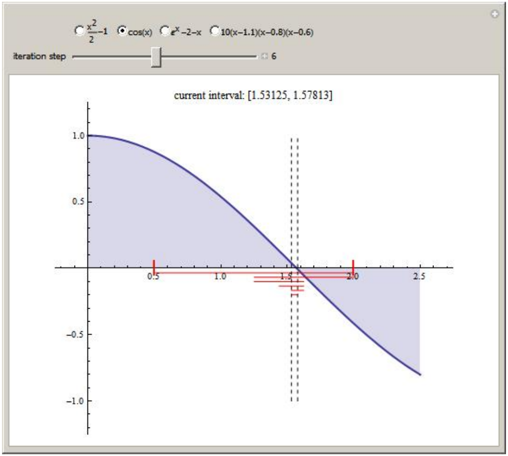
¶¶ Una cota de l'error absolut és: $\frac{|b-a|}{2^n}$ a la n-èsima iteració.

¶¶ Per a aplicar el mètode considerem 3 successions: $a_n \le p_n \le b_n$ definides per les següents relacions:

$$p_n = \frac{a_n+b_n}{2},\ a_{n+1} = \begin{cases} a_n & \text{si } f(a_n)\cdot f(p_n)<0 \\ p_n & \text{si } f(a_n)\cdot f(p_n)>0 \end{cases},\ b_{n+1} = \begin{cases} p_n & \text{si } f(a_n)\cdot f(p_n)<0 \\ b_n & \text{si } f(a_n)\cdot f(p_n)>0 \end{cases}$$

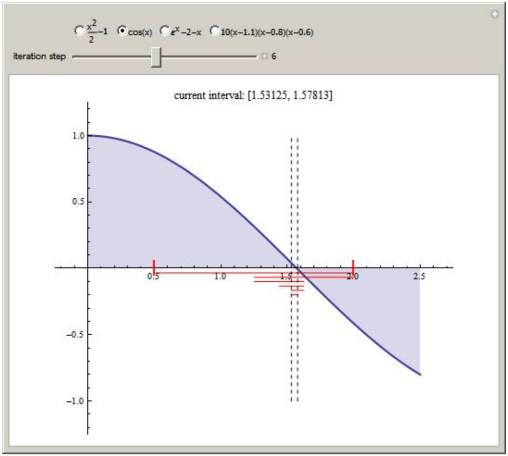
On els valors inicials vénen donats per: $a_0 := a, \quad b_0 := b$

Mètode de la bisecció: $f(x) = \cos x$ (figura 1a)



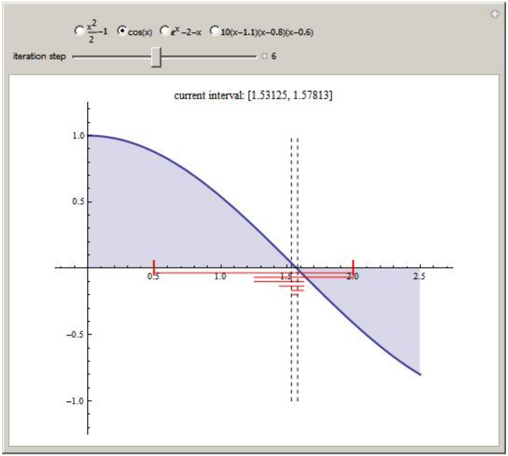
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
1	0.5	2.0	1.25	0.87758256189	-0.416146836547	0.315322362395

Mètode de la bisecció: $f(x) = \cos x$ (figura 1a)



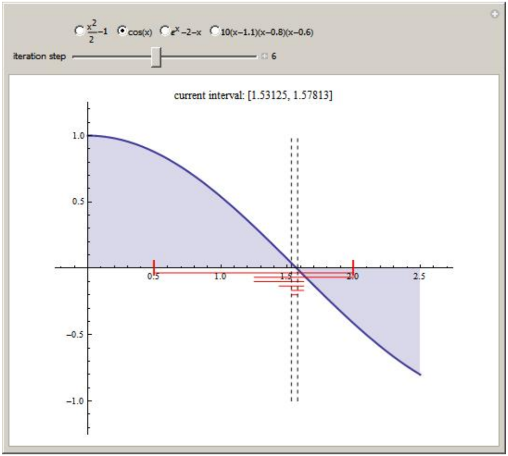
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
1	0.5	2.0	1.25	0.87758256189	-0.416146836547	0.315322362395
2	1.25	2.0	1.625	0.315322362395	-0.416146836547	-0.054177135027

Mètode de la bisecció: $f(x) = \cos x$ (figura 1a)



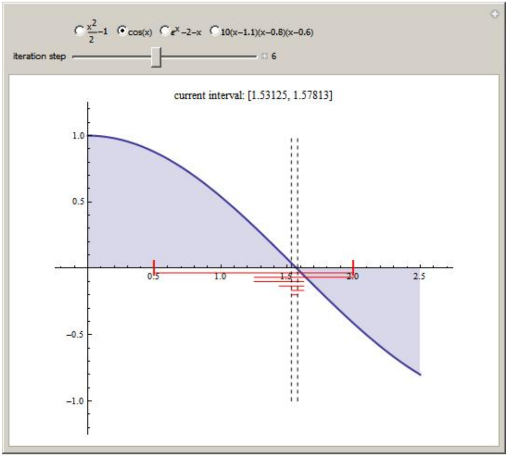
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
1	0.5	2.0	1.25	0.87758256189	-0.416146836547	0.315322362395
2	1.25	2.0	1.625	0.315322362395	-0.416146836547	-0.054177135027
3	1.25	1.625	1.4375	0.315322362395	-0.054177135027	0.132901944453

Mètode de la bisecció: $f(x) = \cos x$ (figura 1a)



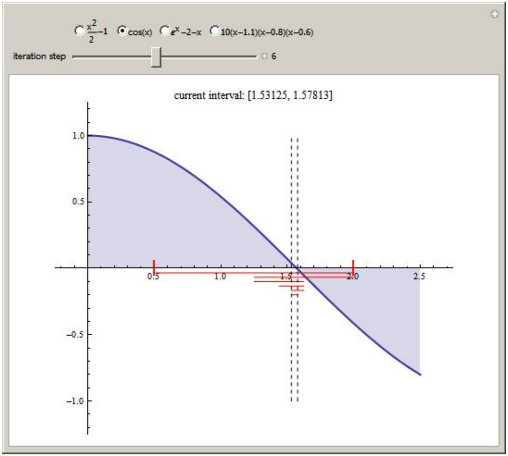
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
1	0.5	2.0	1.25	0.87758256189	-0.416146836547	0.315322362395
2	1.25	2.0	1.625	0.315322362395	-0.416146836547	-0.054177135027
3	1.25	1.625	1.4375	0.315322362395	-0.054177135027	0.132901944453
4	1.4375	1.625	1.53125	0.132901944453	-0.054177135027	0.039536019772

Mètode de la bisecció: $f(x) = \cos x$ (figura 1a)



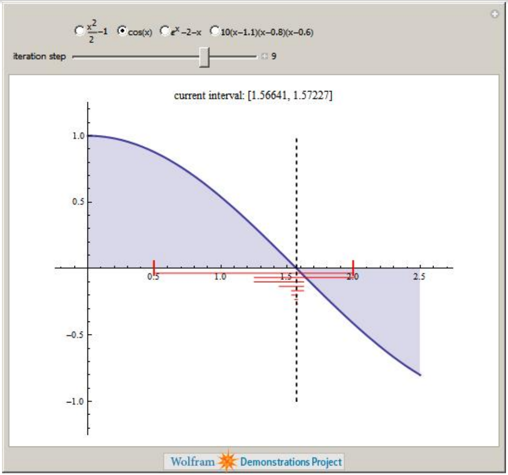
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
1	0.5	2.0	1.25	0.87758256189	-0.416146836547	0.315322362395
2	1.25	2.0	1.625	0.315322362395	-0.416146836547	-0.054177135027
3	1.25	1.625	1.4375	0.315322362395	-0.054177135027	0.132901944453
4	1.4375	1.625	1.53125	0.132901944453	-0.054177135027	0.039536019772
5	1.53125	1.625	1.578125	0.039536019772	-0.054177135027	-0.007328607602

Mètode de la bisecció: $f(x) = \cos x$ (figura 1a)



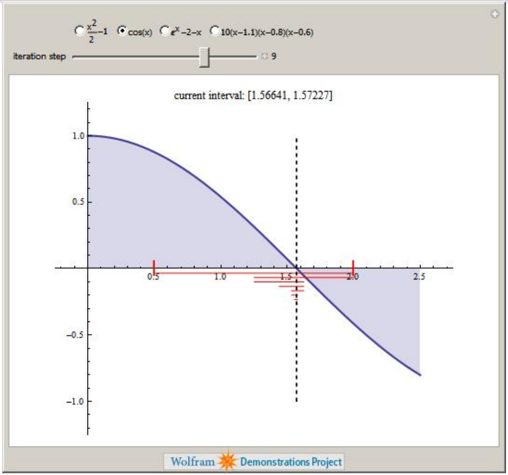
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
1	0.5	2.0	1.25	0.87758256189	-0.416146836547	0.315322362395
2	1.25	2.0	1.625	0.315322362395	-0.416146836547	-0.054177135027
3	1.25	1.625	1.4375	0.315322362395	-0.054177135027	0.132901944453
4	1.4375	1.625	1.53125	0.132901944453	-0.054177135027	0.039536019772
5	1.53125	1.625	1.578125	0.039536019772	-0.054177135027	-0.007328607602
6	1.53125	1.578125	1.5546875	0.039536019772	-0.007328607602	0.016108130112

Mètode de la bisección: $f(x) = \cos x$ (figura 1b)



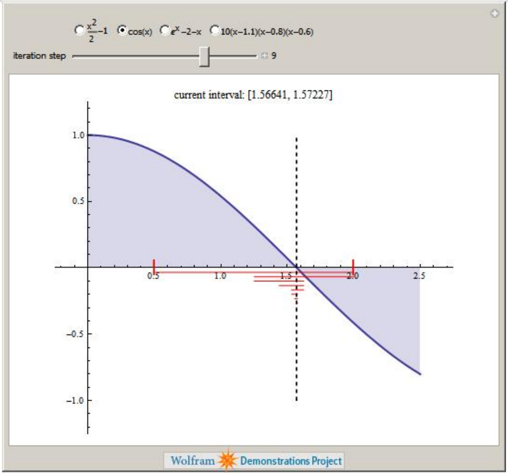
n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
7	1.5546875	1.578125	1.56640625	0.016108130112	-0.007328607602	0.004390062693

Mètode de la bisección: $f(x) = \cos x$ (figura 1b)



n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
7	1.5546875	1.578125	1.56640625	0.016108130112	-0.007328607602	0.004390062693
8	1.56640625	1.578125	1.572265625	0.004390062693	-0.007328607602	-0.001469297676

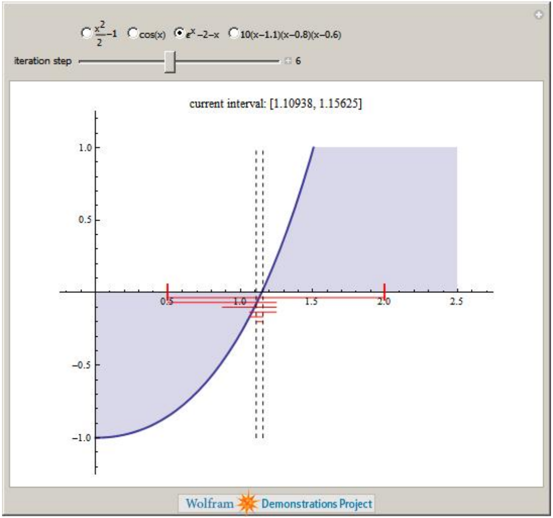
Mètode de la bisección: $f(x) = \cos x$ (figura 1b)



n	a_n	b_n	p_n	$\cos(a_n)$	$\cos(b_n)$	$\cos(p_n)$
7	1.5546875	1.578125	1.56640625	0.016108130112	-0.007328607602	0.004390062693
8	1.56640625	1.578125	1.572265625	0.004390062693	-0.007328607602	-0.001469297676
9	1.56640625	1.572265625	1.57080078125	0.004390062693	-0.001469297676	-0.000004454455

L'arrel de $f(x) = \cos x$ en l'interval $[0.5, 2.5]$ és $x = \pi/2 = 1.5707963267949\ldots$

Mètode de la bisecció: $f(x) = e^x - 2 - x$ (figura 2)



Exercici 1: Confeccionar una taula com la de l'exemple anterior per a la funció $f(x) = e^x - 2 - x$ (vegeu figura), prenent com a interval inicial $[0.5, 2]$.

Mètode de la bisecció: Observacions generals

- Cal notar que, a mesura que anem obtenint noves iteracions, el nombre de xifres significatives que anem deixant inalterades en la successió p_n creix.

Mètode de la bisecció: Observacions generals

- Cal notar que, a mesura que anem obtenint noves iteracions, el nombre de xifres significatives que anem deixant inalterades en la successió p_n creix.

- Hem de tenir en compte que numèricament es treballa en un sistema de punt flotant. Quan $n \gg 1$, a causa dels errors d'arrodoniment, el càlcul del punt mitjà usant l'expressió $p_n = (a_n + b_n)/2$ pot arribar a donar un resultat que estiga fora de l'interval $[a_n, b_n]$. És millor fer servir l'expressió

$$p_n = a_n + \frac{b_n - a_n}{2}$$

Mètode de la bisecció: Observacions generals

- Cal notar que, a mesura que anem obtenint noves iteracions, el nombre de xifres significatives que anem deixant inalterades en la successió p_n creix.

- Hem de tenir en compte que numèricament es treballa en un sistema de punt flotant. Quan $n \gg 1$, a causa dels errors d'arrodoniment, el càlcul del punt mitjà usant l'expressió $p_n = (a_n + b_n)/2$ pot arribar a donar un resultat que estiga fora de l'interval $[a_n, b_n]$. És millor fer servir l'expressió

$$p_n = a_n + \frac{b_n - a_n}{2}$$

- Com que en un sistema de punt flotant l'interval $[0, 1]$ és el que està meś ben representat, convé *normalitzar* les funcions i les variables de manera que es busque la solució a l'esmentat interval.

  Exemple:

$$f(x) = x^3 - 10^6 = 0 \longrightarrow \begin{pmatrix} \text{normalitzem} \\ \tilde{f} = f/10^6 \\ \tilde{x} = x/10^2 \end{pmatrix} \longrightarrow \tilde{f}(\tilde{x}) = \tilde{x}^3 - 1 = 0$$

Mètode de la bisecció: Criteris d'aturada

Criteri d'aturada: condició que s'ha de satisfer per a aturar l'algoritme iteratiu. Per a això se selecciona una tolerància $\epsilon > 0$ i es genera una successió $p_1, \ldots, p_N$, fins que es compleix una de les següents condicions:

1. Tolerància en termes absoluts (sobre l'arrel): $|p_N - p_{N-1}| < \epsilon$.
2. Tolerància sobre el *residu* (sobre la funció): $|f(p_N)| < \epsilon$.
3. Tolerància en termes relatius: $\dfrac{|p_N - p_{N-1}|}{|p_N|} < \epsilon, \quad p_N \neq 0$.

Mètode de la bisecció: Criteris d'aturada

Criteri d'aturada: condició que s'ha de satisfer per a aturar l'algoritme iteratiu. Per a això se selecciona una tolerància $\epsilon > 0$ i es genera una successió $p_1, \ldots, p_N$, fins que es compleix una de les següents condicions:

1. Tolerància en termes absoluts (sobre l'arrel): $|p_N - p_{N-1}| < \epsilon$.
2. Tolerància sobre el *residu* (sobre la funció): $|f(p_N)| < \epsilon$.
3. Tolerància en termes relatius: $\dfrac{|p_N - p_{N-1}|}{|p_N|} < \epsilon, \quad p_N \neq 0$.

Cadascuna d'aquestes condicions té els seus propis problemes:

1. Hi ha successions $\{p_n\}_{n=0}^{\infty}$ tals que $|p_n - p_{n-1}| \to 0$, però $p_n \to \infty$ quan $n \to \infty$.

   $\Longrightarrow$ Exemple: $p_n = \sum_{k=1}^{n} \frac{1}{k} \xrightarrow{n\to\infty} \infty$, però $p_n - p_{n-1} \longrightarrow 0$.

2. Hi ha situacions en les quals $f(p_n) \simeq 0$, però p_n és significativament diferent de zero.

   $\Longrightarrow$ Exemple: $f(x) = (x-1)^{10}$, $p = 1$, y $p_n = 1 + 1/n$.

   La funció satisfà $|f(p_n)| < 10^{-3}$ si $n > 1$, però $|p_n - p| = 1/n$

   requereix que $n > 1000$ perquè es complisca $|p_n - p| < 10^{-3}$.

3. És el millor criteri, ja que controlem l'error relatiu.

   $\Longrightarrow$ Una modificació útil d'aquest criteri és $2\dfrac{|p_N - p_{N-1}|}{|p_N| + |p_{N-1}|} < \epsilon$.

Continguts

1 Introducció

2 Mètode de la bisecció

3 Mètode de Newton

Mètode de Newton (o de Newton-Raphson): I

Algorisme de cerca d'arrels que opera linealitzant la funció per la recta tangent en un valor inicial.

Mètode de Newton (o de Newton-Raphson): I

Algorisme de cerca d'arrels que opera linealitzant la funció per la recta tangent en un valor inicial.

¶ La convergència global no està garantida. Convé, però, seleccionar un valor inicial prou a prop de l'arrel buscada.

Mètode de Newton (o de Newton-Raphson): I

Algorisme de cerca d'arrels que opera linealitzant la funció per la recta tangent en un valor inicial.

¶ La convergència global no està garantida. Convé, però, seleccionar un valor inicial prou a prop de l'arrel buscada.

¶ Si la funció presenta múltiples punts d'inflexió o pendents grans en l'entorn de l'arrel, llavors l'algorisme pot divergir (*vegeu l'exemple 3*).

Mètode de Newton (o de Newton-Raphson): I

Algorisme de cerca d'arrels que opera linealitzant la funció per la recta tangent en un valor inicial.

¶ La convergència global no està garantida. Convé, però, seleccionar un valor inicial prou a prop de l'arrel buscada.

¶ Si la funció presenta múltiples punts d'inflexió o pendents grans en l'entorn de l'arrel, llavors l'algorisme pot divergir (*vegeu l'exemple 3*).

¶¶ **Algorisme:** Siga $f : [a, b] \to \mathbb{R}$ una funció derivable definida en l'interval [a, b]. Siga el valor inicial x_0. definim $\forall n \in \mathbb{N}$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2}$$

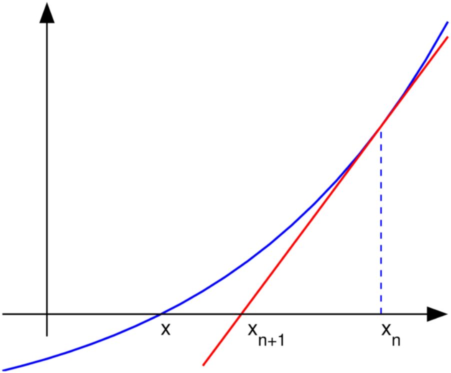
on f' denota la derivada de f.

Mètode de Newton (o de Newton-Raphson): II

¶¶ *Interpretació geomètrica*:

Si per un punt d'iteració tracem la tangent a la corba, el nou punt d'iteració es prendrà com l'abscissa a l'origen de la tangent (punt de tall de la tangent amb l'eix X). Això és equivalent a linealitzar la funció, és a dir, f es reemplaça per una recta tal que conté el punt $(x_0, f(x_0))$ i el pendent coincideix amb la derivada de la funció en el punt, $f'(x_0)$. La nova aproximació a l'arrel, x_1, s'obté a partir de la intersecció de la funció lineal amb l'eix X:

$$f'(x_n) = \frac{f(x_n)}{x_n - x_{n+1}} \qquad (3)$$



La figura il·lustra una iteració del mètode de Newton. Veiem que x_{n+1} és una aproximació millor que x_n per l'arrel x de la funció f.

Mètode de Newton (o de Newton-Raphson): III

¶¶ Una forma alternativa d'obtenir l'algorisme és desenvolupant la funció $f(x)$ en sèrie de Taylor, al voltant del punt x_n:

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + (x - x_n)^2 \frac{f''(x_n)}{2!} + \dots \tag{4}$$

Si x_{n+1} està prou a prop de x_n es pot truncar el desenvolupament a partir del terme d'ordre 2:

$$f(x_{n+1}) = f(x_n) + f'(x_n)(x_{n+1} - x_n) \tag{5}$$

Si a més s'accepta que x_{n+1} tendeix a l'arrel, s'ha de complir que $f(x_{n+1}) = 0$ després, substituint en l'expressió anterior, obtenim l'algorisme.

Mètode de Newton (o de Newton-Raphson): III

¶¶ Una forma alternativa d'obtenir l'algorisme és desenvolupant la funció $f(x)$ en sèrie de Taylor, al voltant del punt x_n:

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + (x - x_n)^2 \frac{f''(x_n)}{2!} + \dots \tag{4}$$

Si x_{n+1} està prou a prop de x_n es pot truncar el desenvolupament a partir del terme d'ordre 2:

$$f(x_{n+1}) = f(x_n) + f'(x_n)(x_{n+1} - x_n) \tag{5}$$

Si a més s'accepta que x_{n+1} tendeix a l'arrel, s'ha de complir que $f(x_{n+1}) = 0$ després, substituint en l'expressió anterior, obtenim l'algorisme.

¶¶ **Convergència del Mètode:** Almenys, **quadràtica**. No obstant això, si l'arrel buscada és de multiplicitat algebraica major a un (c.a., una arrel doble, triple...), el mètode passa a ser lineal.

Mètode de Newton (o de Newton-Raphson): III

¶¶ Una forma alternativa d'obtenir l'algorisme és desenvolupant la funció $f(x)$ en sèrie de Taylor, al voltant del punt x_n:

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + (x - x_n)^2 \frac{f''(x_n)}{2!} + \dots \tag{4}$$

Si x_{n+1} està prou a prop de x_n es pot truncar el desenvolupament a partir del terme d'ordre 2:

$$f(x_{n+1}) = f(x_n) + f'(x_n)(x_{n+1} - x_n) \tag{5}$$

Si a més s'accepta que x_{n+1} tendeix a l'arrel, s'ha de complir que $f(x_{n+1}) = 0$ després, substituint en l'expressió anterior, obtenim l'algorisme.

¶¶ **Convergència del Mètode:** Almenys, **quadràtica**. No obstant això, si l'arrel buscada és de multiplicitat algebraica major a un (c.a., una arrel doble, triple...), el mètode passa a ser lineal.

¶¶ **Estimació de l'Error:** Com que el mètode de Newton convergeix quadràticament, si α és l'arrel, i donada una constant C, es verifica:

$$|x_{k+1} - \alpha| \leq C|x_k - \alpha|^2 \tag{6}$$

Això vol dir que si en algun moment l'error és menor o igual a 0.1, a cada nova iteració dupliquem (aproximadament) el nombre de decimals exactes.

L'error relatiu s'estima emprant els valors entre dues iteracions successives:

$$E = \frac{|x_{k+1} - x_k|}{|x_{k+1}|} \tag{7}$$

Amb la qual cosa es pren l'error relatiu com si l'última aproximació fóra el valor exacte. S'atura el procés iteratiu quan aquest error relatiu és aproximadament menor que una quantitat fixada prèviament.

Mètode de Newton (o de Newton-Raphson): IV

¶¶ Exemple: Trobem un nombre positiu x (amb 11 xifres significatives) tal que $\cos x = x^3$.

Solució: Es tracta de trobar el zero de $f(x) = \cos x - x^3$. Sabem que $f'(x) = -\sin x - 3x^2$. Com que $\cos x \leq 1\ \forall x$ i $x^3 > 1\ \forall x > 1$, deduïm que el zero està entre 0 i 1.

Prenguem el valor inicial: $x_0 = 0.5$

$$\begin{array}{llllll}
x_1 & = & x_0 - \frac{f(x_0)}{f'(x_0)} & = & 0,5 - \frac{\cos(0,5)-0,5^3}{-\sin(0,5)-3\times 0,5^2} & = 1,112141637097 \\
x_2 & = & x_1 - \frac{f(x_1)}{f'(x_1)} & & \vdots & = \underline{0},909672693736 \\
x_3 & & \vdots & & \vdots & = \underline{0,8}67263818209 \\
x_4 & & \vdots & & \vdots & = \underline{0,86547}7135298 \\
x_5 & & \vdots & & \vdots & = \underline{0,8654740331}11 \\
x_6 & & \vdots & & \vdots & = \underline{0,865474033102}
\end{array} \qquad (8)$$

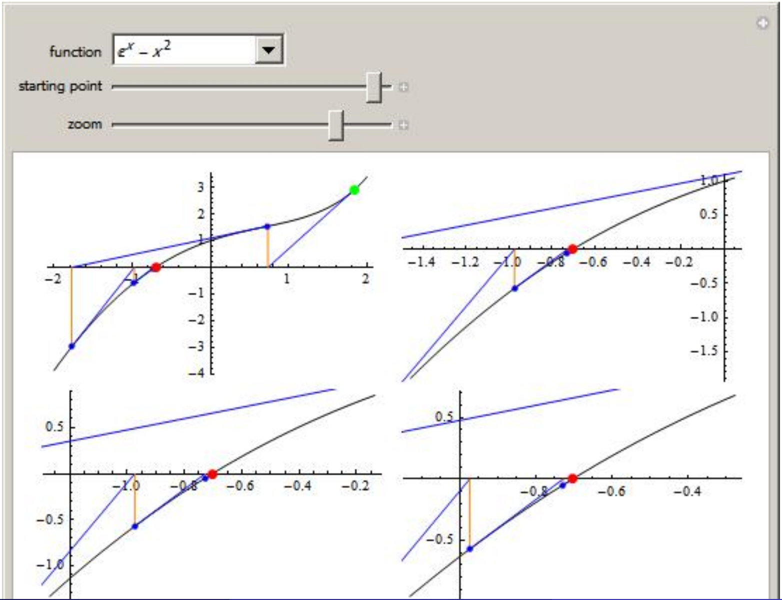
Els dígits correctes estan subratllats. En particular, x_6 és correcte per al nombre de decimals demanats. Podem veure que el nombre de dígits correctes després de la coma s'incrementa des de 2 (per x_3) fins a 5 i 10, il·lustrant la convergència quadràtica.

Exercici 2: Calculeu l'arrel positiva de $\cos x = x^3$ amb 11 xifres significatives partint dels següents valors inicials:

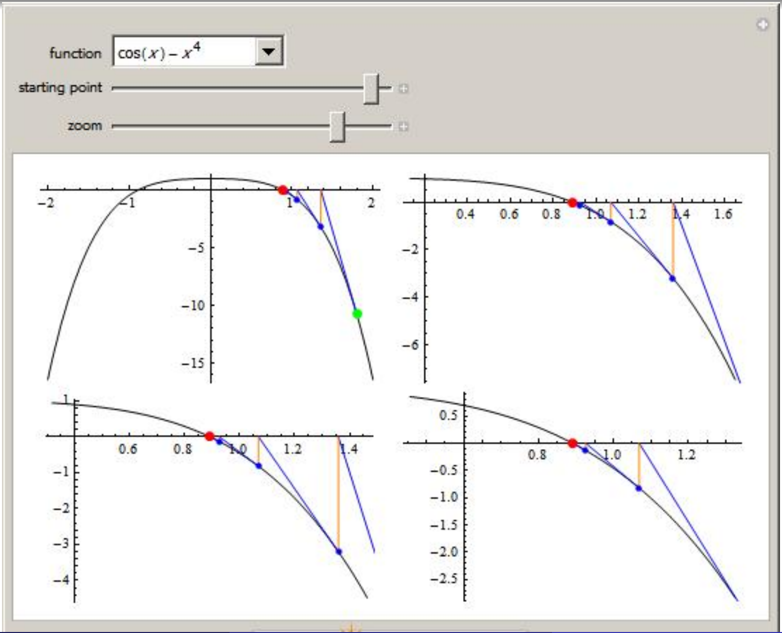
1. $x_0 = 0.1$. Es necessiten més o menys iteracions que en el cas de l'exemple ($x_0 = 0.5$)?
2. $x_0 = 0$. Comenteu el resultat.

Mètode de Newton (o de Newton-Raphson). V: Exemple 1

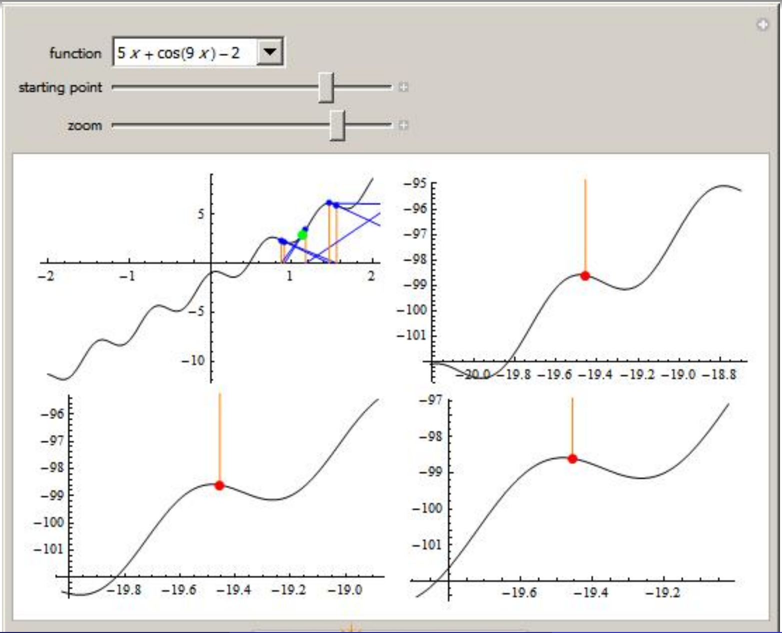
Demo



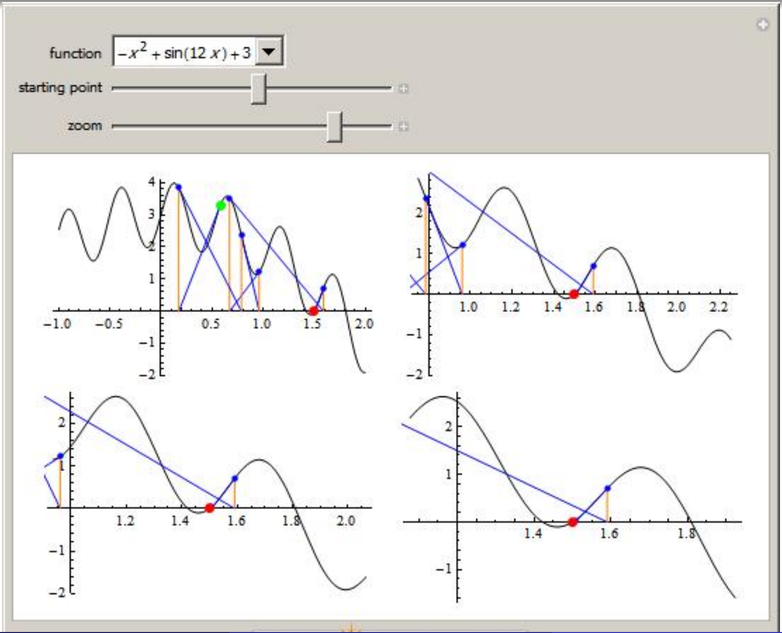
Mètode de Newton (o de Newton-Raphson). VI: Exemple 2



Mètode de Newton-Raphson. VII: Exemple 3



Mètode de Newton-Raphson. VIII: Exemple 4



Comparació de mètodes de cerca d'arrels

	Bisecció	Newton-Raphson
Inici de la cerca	especifiquem interval $[a, b]$	només un punt inicial x_0
Error en cada iteració	$\frac{(b_n - a_n)}{2}$	$\lvert x_n - x_{n-1} \rvert$
Error després n iteracions	$\frac{(b - a)}{2^n}$	Cal avaluar-ho en cada cas
Convergència	Lineal (garantida sii T. de Bolzano)	Quadràtica (no garantida!)

Comparació de mètodes de cerca d'arrels

	Bisecció	Newton-Raphson
Inici de la cerca	especifiquem interval $[a, b]$	només un punt inicial x_0
Error en cada iteració	$\dfrac{(b_n - a_n)}{2}$	$\lvert x_n - x_{n-1} \rvert$
Error després n iteracions	$\dfrac{(b - a)}{2^n}$	Cal avaluar-ho en cada cas
Convergència	Lineal (garantida sii T. de Bolzano)	Quadràtica (no garantida!)

Exercici 3: Utilitzeu els mètodes de Newton-Raphson i de la bisecció per a trobar una arrel real de l'equació $F(x) = x^3 + 2x^2 + 10x - 20 = 0$. L'aproximació a l'arrel ha de tenir una precisió absoluta $\epsilon < 10^{-3}$. Tingueu en compte que la precisió absoluta es defineix de manera diferent en el cas del mètode de Newton-Raphson i en el de la bisecció. Quin dels dos mètodes convergeix més de pressa a la solució?

Mètode de Newton-Raphson: extensió a sistemes

El mètode es pot generalitzar a sistemes d'equacions no lineals

$$\mathbf{f}(\mathbf{x}+\Delta\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x})\Delta\mathbf{x} + \mathcal{O}\left(\frac{\partial^2 f_i}{\partial x_j \partial x_k}\right)$$

on

$$J_{ij} = \frac{\partial f_i}{\partial x_j},$$

és la matriu Jacobiana associada a $\mathbf{f}(\mathbf{x})$.

Si som a prop d'una arrel

$$\mathbf{f}(\mathbf{x}+\Delta\mathbf{x}) \simeq \mathbf{0},$$

i s'arriba a

$$\mathbf{J}(\mathbf{x})\Delta\mathbf{x} \simeq -\mathbf{f}(\mathbf{x}) \tag{9}$$

Procediment: partim d'un vector de valors inicials $\mathbf{x}_0$ i s'aplica l'Eq. 9 d'on obtenim $\Delta\mathbf{x}$. El nou valor de $\mathbf{x}$ es calcula usant

$$\mathbf{x} = \mathbf{x}_0 + \Delta\mathbf{x}. \tag{10}$$

Deprés repetim l'algorisme iterativament.

Mètode de Newton-Raphson: extensió a sistemes

Exemple:

Determineu els punts d'intersecció de les dues funcions:

$$\begin{aligned} x^2 + y^2 &= 3 \\ xy &= 1 \end{aligned}$$

El problema és equivalent a

$$\begin{aligned} f_1(x, y) &= x^2 + y^2 - 3 = 0 \\ f_2(x, y) &= xy - 1 = 0 \end{aligned}$$

$$\mathbf{J}(x, y) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}$$

Apliquem $\mathbf{J}(\mathbf{x})\Delta\mathbf{x} = -\mathbf{f}(\mathbf{x})$

$$\begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} -x^2 - y^2 + 3 \\ -xy + 1 \end{pmatrix} \tag{11}$$

Mètode de Newton-Raphson: extensió a sistemes

1. Iteració: Prenguem com a valors inicials $x_0 = 0.5$ i $y_0 = 1.5$ i resolguem per a trobar $(\Delta x_0, \Delta y_0)$:

$$\begin{pmatrix} 1.0 & 3.0 \\ 1.5 & 0.5 \end{pmatrix} \begin{pmatrix} \Delta x_0 \\ \Delta y_0 \end{pmatrix} = \begin{pmatrix} 0.50 \\ 0.25 \end{pmatrix} \Longrightarrow \begin{pmatrix} \Delta x_0 \\ \Delta y_0 \end{pmatrix} = \begin{pmatrix} 0.125 \\ 0.125 \end{pmatrix}$$

   Calulem els nous valors de x i de y utilitzant l'expressió (10):

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} \Delta x_0 \\ \Delta y_0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix} + \begin{pmatrix} 0.125 \\ 0.125 \end{pmatrix} = \begin{pmatrix} \underline{0.6}25 \\ \underline{1.6}25 \end{pmatrix}$$

2. Iteració: Substituïm en l'Eq. (11) els nous valors: $x = x_1$, $y = y_1$ i calculem $(\Delta x_1, \Delta y_1)$:

$$\begin{pmatrix} 1.250 & 3.250 \\ 1.625 & 0.625 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta y_1 \end{pmatrix} = \begin{pmatrix} -0.031250 \\ -0.015625 \end{pmatrix} \Longrightarrow \begin{pmatrix} \Delta x_1 \\ \Delta y_1 \end{pmatrix} = \begin{pmatrix} -0.0069\bar{4} \\ -0.0069\bar{4} \end{pmatrix}$$

   Calulem els nous valors de x i de y utilitzant l'expressió (10):

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} \Delta x_1 \\ \Delta y_1 \end{pmatrix} = \begin{pmatrix} 0.625 \\ 1.625 \end{pmatrix} + \begin{pmatrix} -0.0069\bar{4} \\ -0.0069\bar{4} \end{pmatrix} = \begin{pmatrix} \underline{0.6180}\bar{5} \\ \underline{1.6180}\bar{5} \end{pmatrix}$$

Mètode de Newton-Raphson: extensió a sistemes

3. Iteració: Prenguem com a valors inicials $x_2 = 0.6180\bar{5}$ i $y_2 = 1.6180\bar{5}$ i resolguem per a trobar $(\Delta x_2, \Delta y_2)$:

$$\begin{pmatrix} 1.236\bar{1} & 3.236\bar{1} \\ 1.6180\bar{5} & 0.6180\bar{5} \end{pmatrix} \begin{pmatrix} \Delta x_2 \\ \Delta y_2 \end{pmatrix} = \begin{pmatrix} -0.9645062 \cdot 10^{-4} \\ -0.4822531 \cdot 10^{-4} \end{pmatrix} \Longrightarrow \begin{pmatrix} \Delta x_2 \\ \Delta y_2 \end{pmatrix} = \begin{pmatrix} -2.157 \cdot 10^{-4} \\ -2.157 \cdot 10^{-4} \end{pmatrix}$$

Calulem els nous valors de x i de y utilitzant l'expressió (10):

$$\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \begin{pmatrix} \Delta x_2 \\ \Delta y_2 \end{pmatrix} = \begin{pmatrix} 0.6180\bar{5} \\ 1.6180\bar{5} \end{pmatrix} + \begin{pmatrix} -2.157 \cdot 10^{-4} \\ -2.157 \cdot 10^{-4} \end{pmatrix} = \begin{pmatrix} \underline{0.618033988}957902 \\ \underline{1.618033988}957902 \end{pmatrix}$$

Noteu que s'han de prendre TOTES les xifres quan calculem; si no, els errors d'arrodoniment poden dominar el resultat!

Si parem en aquest punt el procés iteratiu, haurem aconseguit, en tan sols TRES iteracions, la solució amb 9 xifres significatives darrere de la coma.

Per a pensar:

- Tenint en compte que el mètode de Newton convergeix a segon ordre, si fem una iteració més, quantes xifres significatives obtindríem?
- Podríeu replantejar el problema resolent una única equació no lineal?

Matemàtiques III:
4. Interpolació polinòmica i integració numèrica

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1 Introducció: Aproximació polinomial

2 Interpolació lineal

3 Interpolació polinòmica (I): Elements bàsics

4 Interpolació polinòmica (II): Newton

5 Interpolació polinòmica (III): Lagrange

6 Introducció: Integració numèrica

7 Fórmules de Newton-Cotes

Continguts

1. Introducció: Aproximació polinomial
2. Interpolació lineal
3. Interpolació polinòmica (I): Elements bàsics
4. Interpolació polinòmica (II): Newton
5. Interpolació polinòmica (III): Lagrange
6. Introducció: Integració numèrica
7. Fórmules de Newton-Cotes

Motivació. Vegeu Burden & Faires, Cap. 3, Introducció

¶ Aproximació:

- Reemplacem una funció f per una altra més simple $\tilde{f}$ amb un grau desitjat d'exactitud.
- S'empra en integració numèrica, on en lloc de calcular $\int_a^b f(x)dx$ calculem $\int_a^b \tilde{f}(x)dx$, ja que $\tilde{f}$ és més senzilla d'integrar (típicament un polinomi).
- També usada quan només coneixem valors discrets $f_i \equiv f(x_i)$, i es pretén construir una funció contínua $\tilde{f}$ que pot representar la llei empírica subjacent sota el conjunt finit de dades.
- Per què no fem servir polinomis de Taylor?

¶ Integració numèrica:

- En la segona part veurem mètodes per a integrar de forma numèrica expressions que, *a priori*, són molt complicades d'integrar.
- La integració és un procediment essencial, ja que operacions bàsiques en enginyeria, com la **convolució**, es basen en la integració.

Introducció. Vegeu Burden & Faires, Cap. 3, Introducció

¶ Siga la funció $f(x)$ que desitgem aproximar utilitzant la classe de funcions $\{g_i(x)\}\ \ i = 0, 1, 2, ..., m$:

$$f(x) \approx F(x) = a_0\ g_0(x) + a_1\ g_1(x) + ... + a_m\ g_m(x) = \sum_{j=0}^{m} a_j\ g_j(x) \quad (1)$$

Si en (1) les $a_i,\ \ i = 0, 1, ..., m$ són constants $\Longrightarrow$ La (1) és una aproximació lineal de $f(x)$ $\Longrightarrow$ Problema: Criteri per a triar les $a_i,\ \ i = 0, 1, ..., m$

Teorema 1 de Weierstrass:

Siga una funció $f(x)$, contínua en l'interval $[a, b]$. Donat qualsevol $\epsilon > 0$, hi ha un $n \equiv [n(\epsilon)]$ i un polinomi $P_n(x)$ de grau n tals que

$$|f(x) - P_n(x)| < \epsilon\ ,\ \forall\ x \in [a, b] \quad (2)$$

Teorema 2 Weierstrass:

Siga una funció $F(t)$, contínua i periòdica de període 2π. Donat qualsevol $\epsilon > 0$, hi ha un $n \equiv [n(\epsilon)]$ i una suma trigonomètrica

$$S_n(t) = a_0 + \Sigma_{k=1}^{n}\ (\ a_k \cos kt + b_k \sin kt) \quad (3)$$

tals que

$$|F(t) - S_n(x)| < \epsilon\ ,\ \forall\ t \quad (4)$$

Continguts

1 Introducció: Aproximació polinomial

2 Interpolació lineal

3 Interpolació polinòmica (I): Elements bàsics

4 Interpolació polinòmica (II): Newton

5 Interpolació polinòmica (III): Lagrange

6 Introducció: Integració numèrica

7 Fórmules de Newton-Cotes

Interpolació

Objectiu

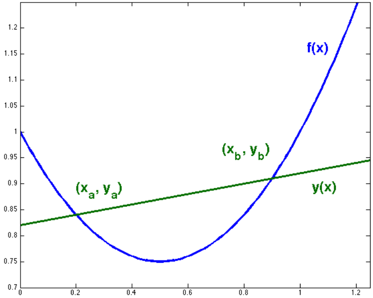
Estimar valors d'una funció en punts no tabulars i, almenys, acotar l'error entre els valors estimats i els exactes.

¶ Siga el conjunt de $m+1$ parells de nombres reals $\{(x_k, y_k)\}_{k=0}^{m}$, tals que $f(x_k) = y_k \ , \ k = 0, 1, \ldots, m$

$\Longrightarrow$ Obtenir la funció F, que anomenarem **funció interpoladora** (associada a f) d'aquests punts. Als punts x_k se'ls anomena **nodes**.

¶ **Interpolació Lineal:** S'utilitzen dos punts, (x_a, y_a) i (x_b, y_b), per a obtenir un tercer punt interpolat (x, y) segons:

$$y = y_a + (x - x_a)\frac{(y_b - y_a)}{(x_b - x_a)} \qquad (5)$$



Continguts

1. Introducció: Aproximació polinomial
2. Interpolació lineal
3. Interpolació polinòmica (I): Elements bàsics
4. Interpolació polinòmica (II): Newton
5. Interpolació polinòmica (III): Lagrange
6. Introducció: Integració numèrica
7. Fórmules de Newton-Cotes

Interpolació polinòmica: Polinomi interpolador i interpolació lineal

Objectiu

Interpolació d'un conjunt de dades o d'una funció per un polinomi (**polinomi interpolador**)

¶¶ Donada una funció f de la qual es coneixen els valors en un nombre finit d'abscisses $x_0, x_1, ..., x_m$, fem a (1) $F(x) := P_n(x)$, és a dir, es tracta de trobar un polinomi $P_n(x)$ (grau: $n \leq m$) tal que:

$$P_n(x_k) \; = \; f(x_k) \; , \; \forall \; k \; = 0, 1, ..., m. \tag{6}$$

al qual s'anomena **polinomi interpolador** de grau n de la funció f.

¶ La interpolació polinòmica és un mètode per a conèixer, d'una manera aproximada, els valors que pren una certa funció de la qual només es coneix la imatge en un nombre finit d'abscisses.

¶ El polinomi interpolador no només ha de satisfer (6) sinó també permetre trobar aproximacions de la funció, en altres punts, dins de la precisió fixada $\Longrightarrow$ Mètodes $\rightarrow$ *polinomi interpolador i fórmula de l'error d'interpolació.*

Continguts

1. Introducció: Aproximació polinomial
2. Interpolació lineal
3. Interpolació polinòmica (I): Elements bàsics
4. Interpolació polinòmica (II): Newton
5. Interpolació polinòmica (III): Lagrange
6. Introducció: Integració numèrica
7. Fórmules de Newton-Cotes

Interpolació de diferències dividides de Newton (I)

Objectiu

Interpolació d'un conjunt de dades, o d'una funció, (x_i, f_i) , $(i = 0, 1, \ldots, m)$
pel **polinomi interpolador** *de grau* $n \leq m$*:*

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \ldots + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1}) \qquad (7)$$

¶ Siguen els $m + 1$ parells de variables discretes (x_i, f_i) , $(i = 0, 1, \ldots, m)$, definits pel punt del domini (abscissa; x_i) de cada una de les dades i la imatge (ordenada; $f(x_i)$) corresponent.

$$f : \ x \to f(x) \ , \ \Longrightarrow \ f(x_i) = f_i, \quad i = 0, 1, \ldots, m \qquad (8)$$

¶¶ Busquem un *polinomi interpolador de grau n ($n \leq m$)* de la forma (7), e.d.:

$$P_n(x) \ = \ \sum_{j=0}^{n} a_j \ p_j(x) \qquad (9)$$

En (1) $\Longrightarrow$ $F(x) := P_n(x)$, $g_j(x) = p_j(x)$, on els $p_j(x)$ són polinomis de grau $j \leq n$:

$$p_0 = 1 \ , \ \ p_{k+1}(x) = \prod_{i=0}^{k}(x - x_i) \ \ \forall \ k = 0, \ldots, n - 1 \qquad (10)$$

Interpolació de diferències dividides de Newton (II): obtenció dels coeficients

BF, §3.2

¶¶ Algorisme: Obtinguem els coeficients a_j.

Calculem el valor del polinomi (7) en els punts tabulars $\{x_i \ , \ i = 0, 1, \ldots, n\}$

$$P_n(x_0) = f(x_0) = a_0 := f[x_0] \tag{11}$$

$\Longrightarrow$ Diferència dividida zero de la funció f respecte a x_i:

$$f[x_i] = f(x_i) \tag{12}$$

$$P_n(x_1) = f(x_1) = a_0 + a_1(x_1 - x_0) \implies a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f[x_1] - f[x_0]}{x_1 - x_0} := f[x_0, x_1] \tag{13}$$

$\Longrightarrow$ Primera diferència dividida de la funció f respecte a x_i y x_{i+1}:

$$f[x_i, x_{i+1}] := \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i} \tag{14}$$

¶¶ De manera inductiva s'obtenen les restants diferències dividides $\Longrightarrow$ Segona diferència dividida de la funció f respecte a x_i, x_{i+1} i x_{i+2}:

$$f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i} \tag{15}$$

Anàlogament, un cop determinades les primeres $(k-1)$ diferències dividides, la k-èsima diferència dividida relativa a $x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+k}$ és:

$$f[x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+k}] = \frac{f[x_{i+1}, x_{i+2}, \ldots, x_{i+k}] - f[x_i, x_{i+1}, \ldots, x_{i+k-1}]}{x_{i+k} - x_i} \tag{16}$$

Com s'indueix a partir de l'obtenció de a_0 i a_1 (11, 13), les constants a_k són:

$$a_k = f[x_0, x_1, x_2, \ldots, x_k] \quad , \quad \forall\ k = 0, 1, 2, \ldots, n \tag{17}$$

Fórmula de diferències dividides interpoladores de Newton:

$$P_n(x) = f[x_0] + \sum_{k=1}^{n} f[x_0, x_1, x_2, \ldots, x_k]\,(x - x_0)(x - x_1)\ldots(x - x_{k-1}) \tag{18}$$

Interpolació de diferències dividides de Newton (III): Cost computacional

Comptem operacions:

$$p_j(x) = (x - x_0)(x - x_1)\dots(x - x_{j-1}) = \underbrace{\prod_{i=0}^{j-1}(x - x_i)}_{(j)\ \text{restes} + (j)\ \text{productes}} \quad \forall\ j = 0, \dots, n-1 \quad (19)$$

És a dir, hem de fer $2j$ operacions per a avaluar p_j. Per a avaluar un terme $a_j p_j(x)$ del sumatori

$$P_n(x) \ = \ \sum_{j=0}^{n} a_j \ p_j(x) \quad (20)$$

hem de fer $2j + 1$ operacions. El nombre total d'operacions serà

$$\sum_{j=0}^{n}(2j+1) = 2\sum_{j=0}^{n} j + (n+1) = 2\frac{(n+0)(n+1)}{2} + (n+1) = (n+1)^2 \sim \mathcal{O}(n^2) \quad (21)$$

Interpolació de diferències dividides de Newton (IV): Exemple

Exemple: Siguen els valors d'una certa funció en els punts tabulars següents:

x_i	1.0	1.3	1.6	1.9	2.2
$f(x_i)$	0.7651977	0.6200860	0.4554022	0.2818186	0.1103623

Es demana: 1) Construir el polinomi interpolador a partir de la fórmula de diferències dividides interpoladores de Newton (18). 2) Interpolar la funció en el punt $x = 1.5$

Solució: 1) Construïm les successives diferències dividides (fins a la quarta):

i	x_i	$f[x_i]$	$f[x_{i-1}, x_i]$	$f[x_{i-2}, x_{i-1}, x_i]$	$f[x_{i-3}, x_{i-2}, \ldots, x_i]$	$f[x_{i-4}, x_{i-3}, \ldots, x_i]$
0	1.0	0.7651977				
			−0.4837057			
1	1.3	0.6200860		−0.1087339		
			−0.5489460		0.0658784	
2	1.6	0.4554022		−0.0494433		0.0018251
			−0.5786120		0.0680685	
3	1.9	0.2818186		0.0118183		
			−0.5715210			
4	2.2	0.1103623				

El polinomi interpolador (18) és:

$$\begin{aligned} P_4(x) \quad = \quad & 0.7651977 - 0.4837057(x-1.0) - 0.1087339(x-1.0)(x-1.3) \\ & +0.0658784(x-1.0)(x-1.3)(x-1.6) + 0.0018251(x-1.0)(x-1.3)(x-1.6)(x-1.9) \end{aligned}$$

2) $P_4\ (1.5) \ = \ 0.5118200$

Continguts

1 Introducció: Aproximació polinomial

2 Interpolació lineal

3 Interpolació polinòmica (I): Elements bàsics

4 Interpolació polinòmica (II): Newton

5 Interpolació polinòmica (III): Lagrange

6 Introducció: Integració numèrica

7 Fórmules de Newton-Cotes

Interpolació de Lagrange

¶¶ Siga f la funció a interpolar, siguen $x_0, x_1, \dots, x_m$ punts tabulars, i siguen $f_0, f_1, \dots, f_m$ les corresponents imatges. **El polinomi interpolador de Lagrange** (de grau n) és

$$L_n(x) \;=\; \sum_{j=0}^{n} f_j l_j(x),\; n \leqslant m \tag{22}$$

on $l_j(x)$ són els **polinomis de Lagrange**:

$$l_j\,(x) = \frac{(x-x_0)\,(x-x_1)\dots(x-x_{j-1})\,(x-x_{j+1})\dots(x-x_n)}{(x_j-x_0)\,(x_j-x_1)\dots(x_j-x_{j-1})\,(x_j-x_{j+1})\dots(x_j-x_n)} \tag{23}$$

¶ Els coeficients $l_j\,(x)$ estan ben definits i són sempre diferents de zero (si $x \neq x_i$).

¶ Cas particular: Si $n = 1 \implies$ *Interpolació lineal*: $L_1(x) \;=\; f(x_0)\,\dfrac{x-x_1}{x_0-x_1} \;+\; f(x_1)\,\dfrac{x-x_0}{x_1-x_0}$

¶ Els coeficients $l_j\,(x)$ es poden reescriure:

$$l_j\,(x) \;=\; \frac{p_{n+1}(x)}{(x-x_j)\;p'_{n+1}(x_j)} \quad,\quad p'_{n+1}(x_j) \;:=\; \left(\frac{d\,p_{n+1}}{d\,x}\right)_{x=xj} \tag{24}$$

$$p_{n+1}(x) \;=\; (x-x_0)(x-x_1)\dots(x-x_n) \tag{25}$$

Error d'interpolació

Siguen $\{x_i\}_{i=0,\dots,m}$ i $I = [x_0, x_m]$. Suposem que la funció f és $n+1$ ($n \leq m$) vegades diferenciable $\forall x \in I$, i que $\xi(x) \in I$.

$$E_n(x) = f(x) - L_n(x) = \left(\frac{p_{n+1}(x)}{(n+1)!}\right)\; f^{(n+1)}(\xi) \tag{26}$$

Interpolació de Lagrange: Cost computacional

Comptem operacions:

$$l_j(x) = \frac{\overbrace{(x-x_0)(x-x_1)\ldots(x-x_{j-1})(x-x_{j+1})\ldots(x-x_n)}^{n\,\text{restes} + n\,\text{productes}}}{\underbrace{(x_j-x_0)(x_j-x_1)\ldots(x_j-x_{j-1})(x_j-x_{j+1})\ldots(x_j-x_n)}_{n\,\text{restes} + n\,\text{productes}}} \tag{27}$$

És a dir, hem de fer $4n+1$ operacions per a avaluar $l_j(x)$. Per a avaluar un terme $f_j l_j(x)$ del sumatori

$$L_n(x) \;=\; \sum_{j=0}^{n} f_j\; l_j(x) \tag{28}$$

hem de fer $4n+2$ operacions. El nombre total d'operacions serà

$$\sum_{j=0}^{n}(4n+2) = (4n+2)\sum_{j=0}^{n} 1 = (4n+2)(n+1) \sim \mathcal{O}(4n^2) \tag{29}$$

Interpolació de Lagrange: Exemple 1

¶ Trobem el valor de la funció $f(x) = \ln x$ a $x = 0.6$ usant un polinomi interpolador de Lagrange de grau 3.

Dades: $f(0.4) = -0.916291$, $f(0.5) = -0.693147$, $f(0.7) = -0.356675$, $f(0.8) = -0.223144$

Solució

Primer.- A partir de (23) amb $x_0 = 0.4$, $x_1 = 0.5$, $x_2 = 0.7$, $x_3 = 0.8$ i les dades anteriors per f_0, f_1, f_2, f_3:

$$l_0(0.6) = -1/6 \ , \ l_1(0.6) = 2/3 \ , \ l_2(0.6) = 2/3 \ , \ l_3(0.6) = -1/6 \tag{30}$$

Segon.- A partir de (22): $L_3(0.60) \approx -0.509975$ a comparar amb el valor "exacte" -0.510826

Tercer.- A partir de (26):

$$E(0.60) = \left(\frac{p_4(0.60)}{(4)!}\right) \left(\frac{-6}{\xi^4}\right) = -\frac{0.0004}{4\,\xi^4} = -\frac{10^{-4}}{\xi^4} \tag{31}$$

En l'interval $[0.4, 0.8]$ es compleix:

$$\frac{10^4}{4096} < \frac{1}{\xi^4} < \frac{10^4}{256} \Longrightarrow \frac{1}{4096} < |E(0.60)| < \frac{1}{256}$$

Interpolació de Lagrange: Exemple 2

¶ Trobem el valor de la funció $f(x) = e^{x+1}$ a $x = 0.75$ usant un polinomi interpolador de Lagrange de grau 2.

Dades: $f(0) = e \;\; , \;\; f\left(\frac{1}{2}\right) = e^{\frac{3}{2}} \;\; , \;\; f(1) = e^2$

Solució: *Primer*.- Els polinomis de Lagrange són:

$$l_0(x) = \frac{\left(x-\frac{1}{2}\right)(x-1)}{\frac{1}{2}} = 2x^2 - 3x + 1 \;\; , \;\; l_1(x) = \frac{x(x-1)}{-\frac{1}{4}} = -4x^2 + 4x \;\; , \;\; l_2(x) = \frac{x\left(x-\frac{1}{2}\right)}{\frac{1}{2}} = 2x^2 - x$$

Segon.- El polinomi interpolador de Lagrange, de grau 2, és:

$$L_2(x) = \sum_{j=0}^{2} f_j l_j(x) = (2e - 4e^{\frac{3}{2}} + 2e^2)x^2 + (-3e + 4e^{\frac{3}{2}} - e^2)x + e$$

Tercer.- Avaluem aquest polinomi en $x = 0.75$:

$$f(0.75) = f\left(\frac{3}{4}\right) = e^{\frac{7}{4}} \simeq p_2\left(\frac{3}{4}\right) \simeq 5.792377$$

Quart.- Estimació de l'error:

Siga el següent valor "exacte": $f\left(\frac{3}{4}\right) = e^{\frac{7}{4}} = 5.754602676\ldots$

$$\Longrightarrow \;\; E_a \simeq |5.792377 - 5.754602616| = 0.037774324 \Rightarrow e_r = \frac{0.037774324}{5.754602616} = 0.006564193 \approx 0.66\,\%$$

Cinquè.- Mètode de les diferències dividides de Newton:

i) Construïm una taula de diferències dividides:

$$f[x_0] = e \;\; , \;\; f[x_1] = e^{\frac{3}{2}} \;\; , \;\; f[x_2] = e^2 \;\; , \;\; f[x_0, x_1] = 2(e^{\frac{3}{2}} - e) \;\; , \;\; f[x_1, x_2] = 2(e^2 - e^{\frac{3}{2}})$$

$$f[x_0, x_1, x_2] = 2(e - 2e^{\frac{3}{2}} + e^2)$$

ii) El polinomi interpolador segons el mètode de Newton és:

$$P_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) = e + 2(e - e^{\frac{3}{2}})x + 2(e^{\frac{3}{2}} - e)x(x - \tfrac{1}{2})$$

$$\Longrightarrow \;\; P_2(x) = (2e - 4e^{\frac{3}{2}} + 2e^2)x^2 + (-3e + 4e^{\frac{3}{2}} - e^2)x + e$$

iii) **S'ha obtingut el mateix polinomi interpolador de Lagrange, amb un menor cost computacional:** $L_2(x) = P_2(x)$

Interpolació de Lagrange: Exemple 2 - Avaluació del polinomi interpolador amb Matlab

¶ Exemple 2: Aprofitarem el resultat obtingut per a avaluar amb Matlab (el polinomi interpolador de grau $n \leq m$ és únic):

```
>> x=[0 1/2 1];                         % introduïm els nodes com un vector
>> fx=inline('exp(x+1)','x');           % declarem la funció usant inline
>> fxv=feval(fx, x);                    % valors de la funció en els nodes
>> Lagrange = polyfit(x, fxv, 2)        % Polinomi interpolador
>> fxi=polyval(Lagrange,0.75)           % Donem valors a 0.75
>> z=linspace(x(1),x(end),100);         % Generem vector per a donar una representació suau al polinomi
>> fxz=polyval(Lagrange, z);            % Valors del polinomi interpolador a ''z''
>> plot(x, fxv, 'o', z, fxz, 0.75, fxi, 'k*', z, exp(z+1), '--')
```

Continguts

1 Introducció: Aproximació polinomial

2 Interpolació lineal

3 Interpolació polinòmica (I): Elements bàsics

4 Interpolació polinòmica (II): Newton

5 Interpolació polinòmica (III): Lagrange

6 Introducció: Integració numèrica

7 Fórmules de Newton-Cotes

Introducció

Objectiu

Integració numèrica (quadratura numèrica, o simplement quadratura) $\Rightarrow$
Aproximació de la integral definida d'una funció

$$\int_a^b f(x)\,dx \tag{32}$$

¶ Existeixen funcions integrables, però la primitiva no es pot calcular analíticament.

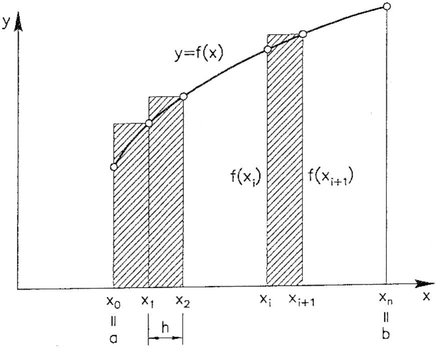
¶ El mètode d'integració més simple descansa en la definició d'integral de Riemann $\Longrightarrow$ Avaluar l'integrant amb increments molt xicotets.

¶ Mètodes basats en funcions d'interpolació: Es basen a aproximar la funció que cal integrar $f(x)$ per una altra $g(x)$ (polinomis, en general) de la qual es coneix la integral exacta, i tal que, en un cert nombre de punts, té el mateix valor que $f(x)$.

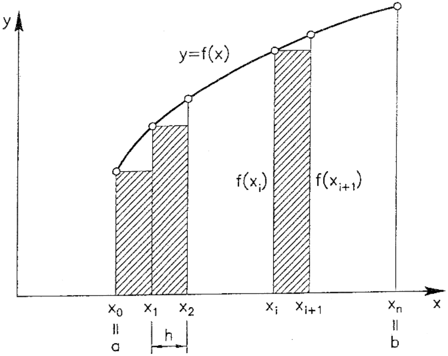
¶ Fórmules de Newton-Cotes: Interpolació amb polinomis, avaluats en punts igualment separats en l'interval $[a, b]$. Exemples: la regla del rectangle, la del trapezi i la de Simpson.

¶ Integrals múltiples: Expressar la integral múltiple com a repetició d'integrals d'una dimensió. $\Longrightarrow$ El cost computacional creix exponencialment amb el nombre de dimensions.

Integral de Riemann: Interpretació geomètrica



Aproximación rectangular superior



Aproximación rectangular inferior

Continguts

1. Introducció: Aproximació polinomial
2. Interpolació lineal
3. Interpolació polinòmica (I): Elements bàsics
4. Interpolació polinòmica (II): Newton
5. Interpolació polinòmica (III): Lagrange
6. Introducció: Integració numèrica
7. Fórmules de Newton-Cotes

Fórmules de Newton-Cotes (I)

Regla del rectangle

La funció interpoladora és una constant (polinomi d'ordre zero) que passa a través del punt $(a,\ f(a))$.

$$\int_a^b f(x)\ dx \ \sim\ (b-a)\ f(a) \tag{33}$$

Regla del punt mitjà

La funció interpoladora és una constant que passa a través del punt $(\frac{a+b}{2},\ f(\frac{a+b}{2}))$

$$\int_a^b f(x)\ dx \ \sim\ (b-a)\ f(\frac{a+b}{2}) \tag{34}$$

Regla del trapezi

La funció interpoladora és una recta (polinomi de grau 1) que passa a través dels punts $(a,\ f(a))$ i $(b,\ f(b))$.

$$\int_a^b f(x)\ dx \ \sim\ (b-a)\ \frac{f(a)+f(b)}{2} \tag{35}$$

Regla de Simpson

La funció interpoladora és un polinomi de grau 2 que passa a través dels punts $(a,\ f(a))$, $(\frac{a+b}{2},\ f(\frac{a+b}{2}))$ i $(b,\ f(b))$.

$$\int_a^b f(x)\,dx \ \approx\ \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right] \tag{36}$$

Fórmules de Newton-Cotes (II). Errors.

Regla del rectangle

La funció interpoladora és un polinomi d'ordre zero $\Rightarrow$ donarà resultats exactes per a funcions constants; si no, tindrem un error de primer ordre:

$$\int_a^b f(x)\ dx = (b-a)\ f(a) + \frac{1}{2} f'(\xi_R)(b-a)^2; \qquad \xi_R \in [a,b] \tag{37}$$

Regla del punt mitjà

La funció interpoladora és un polinomi d'ordre zero $\Rightarrow$ donarà resultats exactes fins a funcions lineals; si no, tindrem un error de segon ordre:

$$\int_a^b f(x)\ dx = (b-a)\ f(\frac{a+b}{2}) + \frac{1}{24} f''(\xi_M)(b-a)^3; \qquad \xi_M \in [a,b] \tag{38}$$

Regla del trapezi

La funció interpoladora és un polinomi d'ordre u $\Rightarrow$ donarà resultats exactes fins a funcions lineals, si no tindrem un error de segon ordre:

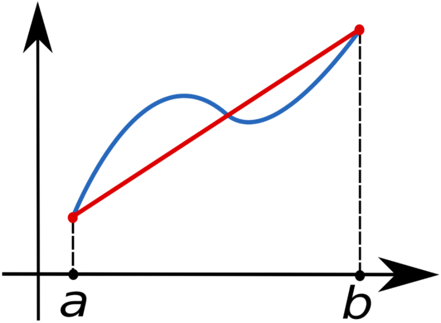
$$\int_a^b f(x)\ dx = (b-a)\ \frac{f(a)+f(b)}{2} + \frac{1}{12} f''(\xi_T)(b-a)^3; \qquad \xi_T \in [a,b] \tag{39}$$

Regla de Simpson

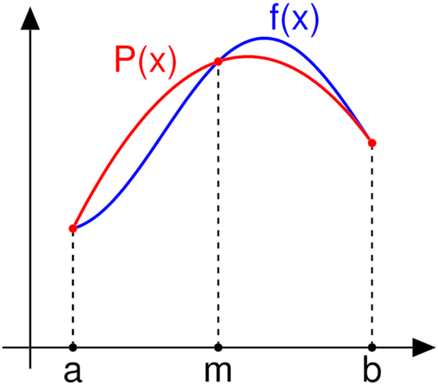
La funció interpoladora és un polinomi de grau 2 $\Rightarrow$ donarà resultats exactes fins a funcions cúbiques; si no, tindrem un error de quart ordre:

$$\int_a^b f(x)\,dx = \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] + \frac{1}{90} f^{(IV)}(\xi_S)(b-a)^5; \qquad \xi_S \in [a,b] \tag{40}$$

Mètode del trapezi i de Simpson: Interpretació geomètrica



Regla simple del Trapezi



Regla simple de Simpson

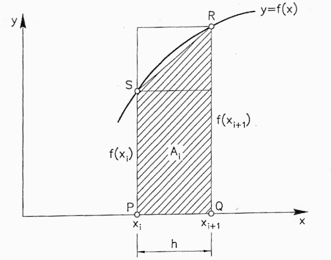
Fórmules **compostes** de Newton-Cotes (I).

La utilitat de les regles simples és prou limitada, a causa de la impossibilitat de controlar l'error en l'aproximació. Les regles compostes es basen en l'aplicació de regles simples sobre malles de punts igualment espaiades

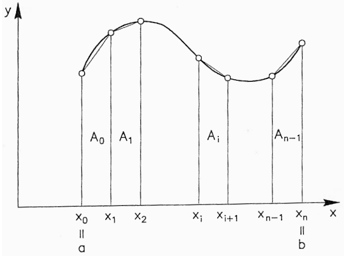
$$a = x_0 < x_1 < \cdots < x_n = b,$$

amb pas $h = (b-a)/N$, que cobreixen l'interval d'integració.

Exemple: Regles simple i composta del trapezi.



Aproximación mediante un trapecio



El método compuesto del trapecio

Fórmules **compostes** de Newton-Cotes (II).

L'error de cada regla composta és proporcional a una potència de h. Així, h es pot determinar amb la condició que l'error siga inferior a una cota ϵ prefixada, de manera que aquestes regles permeten, en la pràctica, un control efectiu sobre l'error de integració. Així, és possible garantir, *a priori*, una certa precisió en l'aproximació obtinguda.

A continuació resumim les regles compostes bàsiques, juntament amb els seus errors associats:

Rectangle	$R_N(f) = h\sum_{i=0}^{N-1} f_i$	$E_{R_N} = \dfrac{h}{2} f'(\xi_R)(b-a)$
Punt Mitjà	$M_N(f) = h\sum_{i=0}^{N-1} f_{i+\frac{1}{2}}$	$E_{M_N} = \dfrac{h^2}{24} f''(\xi_M)(b-a)$
Trapezi	$T_N(f) = \dfrac{h}{2}(f_0 + f_N) + h\displaystyle\sum_{i=1}^{N-1} f_i$	$E_{T_N} = -\dfrac{h^2}{12} f''(\xi_T)(b-a)$
Simpson	$S_N(f) = \dfrac{h}{6}\left(f_0 + f_N + 2\displaystyle\sum_{i=1}^{N-1} f_i + 4\sum_{i=0}^{N-1} f_{i+\frac{1}{2}}\right)$	$E_{S_N} = -\dfrac{h^4}{180} f^{(IV)}(\xi_S)(b-a)$

on $\xi_R, \xi_M, \xi_T, \xi_S$ són punts de l'interval de integració $[a,b]$, desconeguts *a priori*, i $f_i = f(x_i)$, $f_{i+\frac{1}{2}} = f(\frac{x_i + x_{i+1}}{2})$.

Exemples (I)

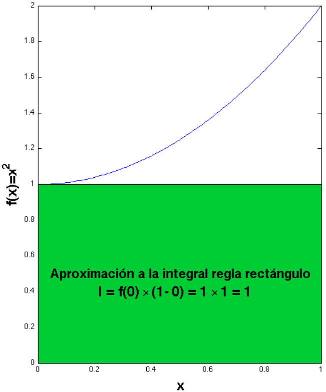
Exemple: *Donada la funció $f(x) = x^2 + 1$, apliqueu les quatre regles bàsiques per a obtenir una aproximació a $I = \int_0^1 f(x)dx$.*

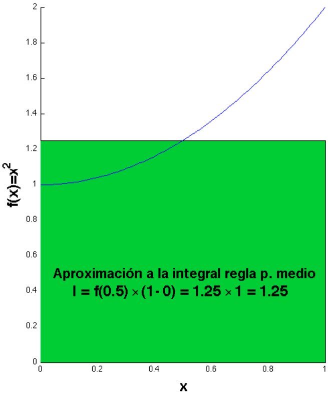
Solució:

1. Identifiquem a, b i $m = (a+b)/2$, els valors dels quals són $0, 1,$ i 0.5, respectivament.
2. Construïm la taula de valors:

x	0	0.5	1
$f(x)$	1	1.25	2

3. Calculem les integrals amb les fórmules corresponents:





Exemples (II)

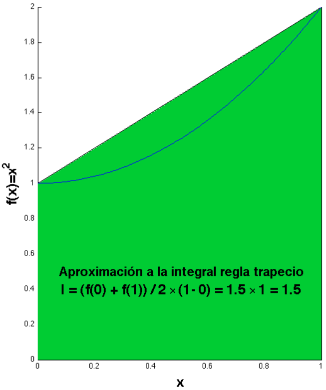
Exemple: *Donada la funció $f(x) = x^2 + 1$, apliqueu les quatre regles bàsiques per a obtenir una aproximació a $I = \int_0^1 f(x)dx$.*

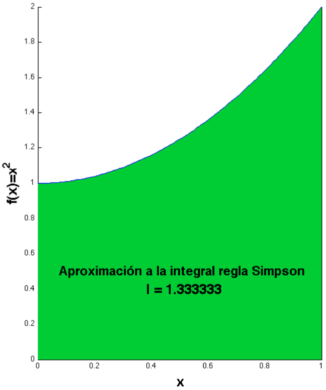
Solució:

1. Identifiquem a, b i $m = (a + b)/2$, els valors dels quals són $0, 1,$ i 0.5, respectivament.
2. Construïm la taula de valors:

x	0	0.5	1
$f(x)$	1	1.25	2

3. Calculem les integrals amb les fórmules corresponents:





Exercicis

Exercici 1: *Donada la funció* $f(x) = x^3$*, apliqueu les quatre regles bàsiques per a obtenir una aproximació a* $I = \int_1^2 f(x)dx$*. Calculeu també la integral de forma analítica i computeu els errors absoluts i relatius existents en emprar cadascuna de les regles bàsiques. Avalueu els termes d'error corresponents a cadascuna de les fórmules.*

Exercici 2: *Repetiu l'exercici anterior, però considereu ara un interval més gran, en concret, calculeu* $I = \int_1^{10} f(x)dx$*. Els errors absoluts i relatius, augmenten o disminueixen quan l'interval creix?*

Exercici 3: *Repetiu l'exercici anterior, però usant les fórmules de Newton-Cotes compostes amb quatre punts. Estimeu els errors corresponents a cadascuna de les fórmules.*

Matemàtiques III:
5. Mètodes numèrics per a la resolució d'equacions diferencials ordinàries

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1 Introducció

2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

Continguts

1 Introducció

2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

Introducció

Burden & Faires, Cap. 5

Motivació

¶ Tot sistema físic, mecànic, elèctric... sol evolucionar amb el temps d'acord amb determinades expressions que es poden expressar en forma d'equacions que relacionen diferents derivades: equacions diferencials. Quan depenen de només una variable s'anomenen equacions diferencials ordinàries (EDO). Si depenen de diverses variables s'anomenen equacions en derivades parcials.

¶ Qualsevol simulació realista d'un sistema dinàmic utilitzarà la resolució numèrica d'EDOs (ex. evolució temporal de la intensitat en un circuit).

Objectiu

Resoldre **numèricament** equacions diferencials ordinàries explícites, d'ordre n:

$$y^{(n)} = F(x, y, y', \ldots, y^{(n-1)}) \tag{1}$$

on y és una funció desconeguda de x, $y^{(n)}$ és la n-èsima derivada de y, quan la funció y i les seues derivades satisfan n condicions inicials o de contorn:
$y(0) = y_0,\ y'(0) = y'_0, \ldots,\ y^{(n)}(0) = y_0^{(n)}$.

¶ Solució general de (1): $\Longrightarrow$ Una funció n vegades derivable en un interval I de $\mathbb{R}$, que satisfà l'Eq.(1), i que conté n variables arbitràries, corresponents a n constants d'integració.

¶ Solució concreta de (1): $\Longrightarrow$ És l'obtinguda a partir de la solució general mitjançant la fixació de valors particulars per les constants segons siguen les condicions inicials.

Introducció. Exemple 1

¶ Equació de moviment d'una partícula.

$$y'' \equiv \frac{d^2y}{dt^2} = a, \quad y(0) = 0, \quad y'(0) = 1, \tag{2}$$

on a és l'acceleració de la partícula. Si suposem que $a = 1 = constant$, la solució de l'equació diferencial (2) requereix calcular la *solució general* (part homogènia) i una *solució particular*.

Introducció. Exemple 1

¶ Equació de moviment d'una partícula.

$$y'' \equiv \frac{d^2y}{dt^2} = a, \quad y(0) = 0, \quad y'(0) = 1, \tag{2}$$

on a és l'acceleració de la partícula. Si suposem que $a = 1 = constant$, la solució de l'equació diferencial (2) requereix calcular la *solució general* (part homogènia) i una *solució particular*.

¶¶ Solució general de (2):

$$y''(x) = 0 \Longrightarrow y(x) = Ax + B, \tag{3}$$

on A i B són constants d'integració a determinar imposant condicions inicials:

$$\begin{aligned} y(0) &= 0 \Longrightarrow 0 = A \times 0 + B \Longrightarrow B = 0, \\ y'(0) &= 1 \Longrightarrow y'(x) = A \qquad \Longrightarrow A = 1, \Longrightarrow y_{\text{gral}}(x) = x \end{aligned}$$

Introducció. Exemple 1

¶ Equació de moviment d'una partícula.

$$y'' \equiv \frac{d^2y}{dt^2} = a, \quad y(0) = 0, \quad y'(0) = 1, \tag{2}$$

on a és l'acceleració de la partícula. Si suposem que $a = 1 = constant$, la solució de l'equació diferencial (2) requereix calcular la *solució general* (part homogènia) i una *solució particular*.

¶¶ Solució general de (2):

$$y''(x) = 0 \Longrightarrow y(x) = Ax + B, \tag{3}$$

on A i B són constants d'integració a determinar imposant condicions inicials:

$$\begin{aligned} y(0) &= 0 \Longrightarrow 0 = A \times 0 + B \Longrightarrow B = 0, \\ y'(0) &= 1 \Longrightarrow y'(x) = A \qquad \Longrightarrow A = 1, \Longrightarrow y_{\rm gral}(x) = x \end{aligned}$$

¶¶ Solució particular de (2):

$$y_{\rm part}(x) = \frac{x^2}{2} \tag{4}$$

Així doncs, la solució de (2) és

$$y(x) = y_{\rm gral}(x) + y_{\rm part}(x) = \frac{x^2}{2} + x.$$

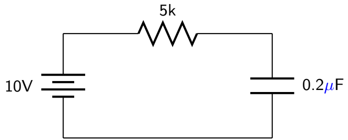
Introducció. Exemple 2

¶ Càlcul de transitoris en un circuit elèctric[Vegeu la pràctica del tema].

Connectem el circuit en un temps $t = 0$. L'objectiu és calcular el corrent després d'un temps $t =$1 ms. El circuit respon a l'EDO:

$$\frac{dI}{dt} + \frac{I}{RC} = \frac{1}{R}\frac{dV}{dt},$$

C=capacitat, R=resistència, I=intensitat, V=tensió



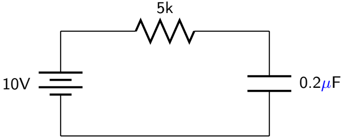
Introducció. Exemple 2

¶ Càlcul de transitoris en un circuit elèctric[Vegeu la pràctica del tema].

Connectem el circuit en un temps $t=0$. L'objectiu és calcular el corrent després d'un temps $t=$1 ms. El circuit respon a l'EDO:

$$\frac{dI}{dt}+\frac{I}{RC}=\frac{1}{R}\frac{dV}{dt},$$

C=capacitat, R=resistència, I=intensitat, V=tensió



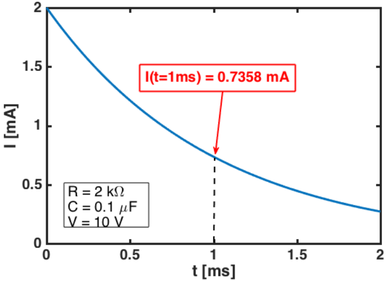
- Si suposem que la tensió no varia amb el temps ($\frac{dV}{dt}=0$) l'EDO anterior se simplifica:

$$\frac{dI}{dt}=-\frac{I}{5\times 0.2}\Longrightarrow\frac{dI}{dt}=-I,$$

- La darrera equació es pot integrar fàcilment:

$$\int\frac{dI}{I}=\int -dt\Rightarrow \log I=-t+\mathcal{K}\Rightarrow I(t)=\mathcal{K}'e^{-t}$$

- La constant d'integració $\mathcal{K}'$ la fixem considerant que quan $t=0$ el condensador està descarregat i, per tant, $I(0)=V/R=2\,$mA $\Rightarrow I(t)=2e^{-t}$.



Continguts

1 Introducció

2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

Diferenciació numèrica (I) Quarteroni, et al., §4.1; Burden & Faires, §4.1

Objectiu

Aproximar les derivades d'una funció en un punt substituint les diferencials avaluades en aquest punt per diferències entre punts propers.

Siga $f : [a, b] \to \mathbb{R}$ una funció contínua i diferenciable en $[a, b]$, i siga $\tilde{x} \in (a, b)$. La derivada de f en $\tilde{x}$ és, per definició

$$f'(\tilde{x}) \equiv \lim_{h \to 0} \frac{f(\tilde{x} + h) - f(\tilde{x})}{h} . \tag{5}$$

Diferenciació numèrica (I)

Quarteroni, et al., §4.1; Burden & Faires, §4.1

Objectiu

Aproximar les derivades d'una funció en un punt substituint les diferencials avaluades en aquest punt per diferències entre punts propers.

Siga $f : [a, b] \to \mathbb{R}$ una funció contínua i diferenciable en $[a, b]$, i siga $\tilde{x} \in (a, b)$. La derivada de f en $\tilde{x}$ és, per definició

$$f'(\tilde{x}) \equiv \lim_{h \to 0} \frac{f(\tilde{x} + h) - f(\tilde{x})}{h} \, . \qquad (5)$$

Partint de (5), suposant que $h > 0$ siga prou "xicoteta", podem construir la següent **aproximació de la derivada** (coneguda com a diferència finita progressiva):

$$f'(\tilde{x}) \approx \Delta_+ f(\tilde{x}) \equiv \frac{f(\tilde{x} + h) - f(\tilde{x})}{h} \qquad f'(\tilde{x}) = \Delta_+ f(\tilde{x}) + \text{ERROR} \qquad (6)$$

Error comès (usant Taylor fins a ordre 2):

$$f(\tilde{x} + h) = f(\tilde{x}) + h f'(\tilde{x}) + \frac{h^2}{2} f''(\xi), \quad \xi \in (\tilde{x}, \tilde{x} + h). \qquad (7)$$

Per tant,

$$\Delta_+ f(\tilde{x}) = f'(\tilde{x}) + \frac{h}{2} f''(\xi) \Longrightarrow \mathcal{O}(h) \, . \qquad (8)$$

Diferenciació numèrica (II) Quarteroni, et al., §4.1; Burden & Faires, §4.1

Objectiu

Aproximar les derivades d'una funció en un punt substituint les diferencials avaluades en aquest punt per diferències entre punts propers.

Usant novament Taylor, tenim que:

$$f(\tilde{x}-h) = f(\tilde{x}) - hf'(\tilde{x}) + \frac{h^2}{2}f''(\eta), \quad \eta \in (\tilde{x}-h, \tilde{x}). \tag{9}$$

cosa que ens permet definir la diferència finita regressiva

$$\Delta_- f(\tilde{x}) \equiv \frac{f(\tilde{x}) - f(\tilde{x}-h)}{h}\,, \tag{10}$$

la qual també compleix que

$$\Delta_- f(\tilde{x}) = f'(\tilde{x}) - \frac{h}{2}f''(\eta) \Longrightarrow \mathcal{O}(h)\,. \tag{11}$$

Diferenciació numèrica (III) Quarteroni, et al., §4.1; Burden & Faires, §4.1

Objectiu

Aproximar les derivades d'una funció en un punt substituint les diferencials avaluades en aquest punt per diferències entre punts propers.

Finalment, podem definir la diferència finita centrada mitjançant

$$\Delta f(\tilde{x}) = \frac{f(\tilde{x}+h) - f(\tilde{x}-h)}{2h}\,, \tag{12}$$

Error comès (usant Taylor fins a ordre 3):

$$\begin{aligned}
f(\tilde{x}+h) &= f(\tilde{x}) + hf'(\tilde{x}) + \frac{h^2}{2}f''(\tilde{x}) + \frac{h^3}{6}f'''(\xi_+), \qquad \xi_+ \in (\tilde{x}, \tilde{x}+h),\\
\underline{f(\tilde{x}-h)} &= \underline{f(\tilde{x}) - hf'(\tilde{x}) + \frac{h^2}{2}f''(\tilde{x}) - \frac{h^3}{6}f'''(\xi_-)}, \qquad \xi_- \in (\tilde{x}-h, \tilde{x}),\\
f(\tilde{x}+h) - f(\tilde{x}-h) &= 2hf'(\tilde{x}) + \frac{h^3}{6}[f'''(\xi_+) + f'''(\xi_-)]
\end{aligned}$$

És a dir que

$$\Delta f(\tilde{x}) = f'(\tilde{x}) + \frac{h^2}{12}f'''(\eta) \Longrightarrow \mathcal{O}(h^2)\,. \tag{13}$$

Diferenciació numèrica (VI) Quarteroni, et al., §4.1; Burden & Faires, §4.1

Resum: aproximació de la derivada primera

Diferència progressiva: $$\Delta_+ f(\tilde{x}) \equiv \frac{f(\tilde{x}+h) - f(\tilde{x})}{h} + \mathcal{O}(h)\,,$$

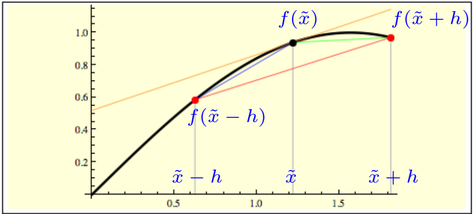
Diferència regressiva: $$\Delta_- f(\tilde{x}) \equiv \frac{f(\tilde{x}) - f(\tilde{x}-h)}{h} + \mathcal{O}(h)\,,$$

Diferència centrada: $$\Delta f(\tilde{x}) \equiv \frac{f(\tilde{x}+h) - f(\tilde{x}-h)}{2h} + \mathcal{O}(h^2)\,.$$

- Les diferències finites solen emprar-se en els anomenats problemes de valors de contorn, on tenim una equació diferencial (p. ex., $y'(x) = f(x,y)$) que ha de satisfer certes condicions de contorn en un interval (a,b) donat. És a dir, es proporcionen els valors (fixos) en els extrems d'aquest interval: $y(a) = y_a$, $y(b) = y_b$.
- En aquest tema no abordarem les tècniques concretes que es fan servir per a resoldre aquests problemes, tot i que es recomana a l'alumne que llija el capítol 11 de Burden & Faires per a complementar els coneixements d'aquesta assignatura.

Diferenciació numèrica (V): Interpretació geomètrica

$$f(x) = \sin x, \quad \tilde{x} = 1.22, \quad h = 0.595$$



Method	Approximation	Absolute true error
forward divided difference	0.0525	0.2912
central divided difference	0.3237	0.0199
backward divided difference	0.5950	0.2513
exact value	0.3436	

Wolfram Demonstrations Project: Finite difference approximations

- Diferència progressiva: Aproxima el pendent de la tangent a $f(x)$ en $\tilde{x}$ pel pendent de la recta que passa pels punts $(\tilde{x}, f(\tilde{x}))$ i $(\tilde{x}+h, f(\tilde{x}+h))$.
- Diferència regressiva: Aproxima el pendent de la tangent a $f(x)$ en $\tilde{x}$ pel pendent de la recta que passa pels punts $(\tilde{x}-h, f(\tilde{x}-h))$ i $(\tilde{x}, f(\tilde{x}))$.
- Diferència centrada: Aproxima el pendent de la tangent a $f(x)$ en $\tilde{x}$ pel pendent de la recta que passa pels punts $(\tilde{x}-h, f(\tilde{x}-h))$ i $(\tilde{x}+h, f(\tilde{x}+h))$.

Diferenciació numèrica (VI)

Diferències finites d'ordre superior:
Podem construir aproximacions a les derivades d'ordre superior usant una definició inductiva. En particular, per a la diferència finita d'ordre 2, tindríem diverses possibilitats

$$\begin{aligned}
\Delta^2_{-} f(\tilde{x}) &= \Delta_{-}(\Delta_{-} f(\tilde{x})) = \Delta_{-}\left(\frac{f(\tilde{x}) - f(\tilde{x}-h)}{h}\right) = \frac{f(\tilde{x}) - 2f(\tilde{x}-h) + f(\tilde{x}-2h)}{h^2} \\
\Delta^2_{+} f(\tilde{x}) &= \Delta_{+}(\Delta_{+} f(\tilde{x})) = \Delta_{+}\left(\frac{f(\tilde{x}+h) - f(\tilde{x})}{h}\right) = \frac{f(\tilde{x}+2h) - 2f(\tilde{x}+h) + f(\tilde{x})}{h^2} \\
\Delta^2_{c+} f(\tilde{x}) &= \Delta(\Delta_{+} f(\tilde{x})) = \Delta\left(\frac{f(\tilde{x}+h) - f(\tilde{x})}{h}\right) = \frac{f(\tilde{x}+2h) - f(\tilde{x}+h) - f(\tilde{x}) + f(\tilde{x}-h)}{2h^2} \\
\Delta^2_{c-} f(\tilde{x}) &= \Delta(\Delta_{-} f(\tilde{x})) = \Delta\left(\frac{f(\tilde{x}) - f(\tilde{x}-h)}{h}\right) = \frac{f(\tilde{x}+h) - f(\tilde{x}) - f(\tilde{x}-h) + f(\tilde{x}-2h)}{2h^2} \\
\Delta^2 f(\tilde{x}) &= \Delta(\Delta f(\tilde{x})) = \Delta\left(\frac{f(\tilde{x}+h) - f(\tilde{x}-h)}{2h}\right) = \frac{f(\tilde{x}+2h) - 2f(\tilde{x}) + f(\tilde{x}-2h)}{4h^2} \\
\Delta^2_{+-} f(\tilde{x}) &= \Delta_{+}(\Delta_{-} f(\tilde{x})) = \Delta_{+}\left(\frac{f(\tilde{x}) - f(\tilde{x}-h)}{h}\right) = \frac{f(\tilde{x}+h) - 2f(\tilde{x}) + f(\tilde{x}-h)}{h^2} \\
\Delta^2_{-+} f(\tilde{x}) &= \Delta_{-}(\Delta_{+} f(\tilde{x})) = \Delta_{-}\left(\frac{f(\tilde{x}+h) - f(\tilde{x})}{h}\right) = \frac{f(\tilde{x}+h) - 2f(\tilde{x}) + f(\tilde{x}-h)}{h^2}
\end{aligned}$$

Se sol prendre $\Delta^2_{+-} f(\tilde{x})$ per ser simètrica i dependre de punts més propers a $\tilde{x}$.

Exercici 1: Comproveu que la diferència finita de segon ordre simètrica reprodueix algunes propietats bàsiques de les derivades segones:

$$\begin{aligned}
f(x) = A &\implies \Delta^2_{+-} f(x) = 0 \\
f(x) = Bx + A &\implies \Delta^2_{+-} f(x) = 0 \\
f(x) = Cx^2 + Bx + A &\implies \Delta^2_{+-} f(x) = 2C
\end{aligned}$$

Continguts

1 Introducció

2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

El problema de Cauchy (I)

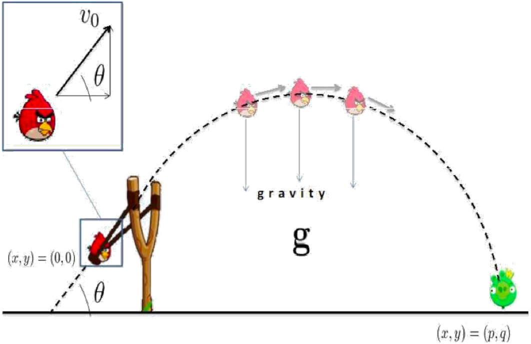
Problema de valors inicials (PVI) o Problema de Cauchy

Trobar $y : I \subset \mathbb{R} \to \mathbb{R}$ tal que

$$\begin{cases} y'(x) = f(x, y(x)) \quad \forall x \in I, & \text{EDO de primer ordre} \\ y(x_0) = y_0, & \text{Dada inicial} \end{cases} \qquad (14)$$

on $f : I \times \mathbb{R} \to \mathbb{R}$ és una funció, $y'(x) \equiv \frac{dy}{dx}$, $x_0 \in I$, i el valor y_0 és la dada inicial.

A.L. Cauchy (1789-1857)



El problema de Cauchy (II)

¶ Es poden resoldre **tots** els problemes de Cauchy?

El problema de Cauchy (II)

¶ Es poden resoldre **tots** els problemes de Cauchy? $\Longrightarrow$ NO, cal que el problema estiga **ben condicionat**!

Resultat clàssic de l'anàlisi

Suposem que la funció $f(x, y(x))$ és:

1. contínua respecte als seus dos arguments;
2. Lipschitziana respecte al seu segon argument, és a dir, existeix una constant positiva L (constant Lipschiziana) tal que

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2| \quad \forall x \in I, \quad \forall y_1, y_2 \in \mathbb{R}.$$

Aleshores el problema està ben condicionat:

- La solució $y = y(x)$ del PVI (14) existeix i és única.
- Una "petita pertorbació" de les equacions (p. e., discretització) produeix una pertorbació "igual de petita" en la solució.

Continguts

1 Introducció

2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

Generalitats

Quan la solució de l'equació diferencial no és analítica, o no és explícita, podem avaluar la solució numèricament.

¶ Estratègia comuna en els mètodes de solució:

1. Se subdivideix l'interval $I = [x_o, x_f]$ en n subintervals equiespaiats $\{x_i\}_{i=0}^{n}$, d'amplària $h = \dfrac{x_f - x_0}{n}$; h és l'anomenat *pas de la discretització*, de forma que:
$$x_i = x_0 + ih, \qquad 0 \leq i \leq n \tag{15}$$
2. En cadascun dels nodes x_i busquem el valor desconegut u_i, el qual s'aproxima a $y_i = y(x_i)$.
3. El conjunt de valors $\{u_0 = y_0, u_1, \ldots, u_n\}$ representa la solució numèrica.

Continguts

1 Introducció

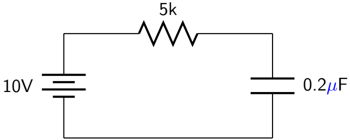
2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x, y),\ y(x_0) = y_0$

Tornem a l'exemple del circuit:



$$\underbrace{\underbrace{\frac{dI}{dt} = -\frac{I}{RC}}_{\text{EDO 1r ordre}} \quad ; \quad \underbrace{I(0) = I_0}_{\text{condició inicial}}}_{\text{PVI}} \quad \Longrightarrow \quad \underbrace{\frac{du}{dx} = -\frac{u}{RC} \quad ; \quad u(x_0) = u_0}_{\text{canvi de notació per conveniència}}$$

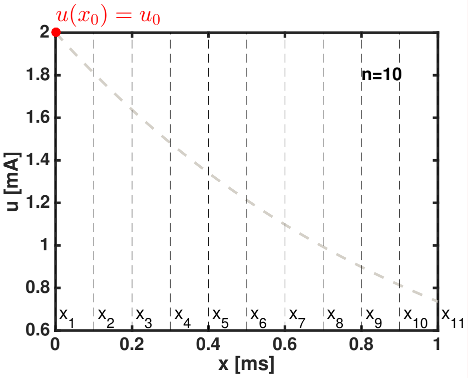
Volem calcular una solució aproximada d'aquest PVI en $t \in [0, 1]$, o equivalentment en $x \in [x_0, x_f]$, on fem les identificacions $x \leftrightarrow t;\ x_0 \leftrightarrow 0;\ x_f \leftrightarrow 0.$

A més, el terme de la dreta de l'EDO és una funció de la variable independent, x, (encara que implícitament) i de la funció $u(x)$, que és la solució que volem trobar numèricament:

$$-\frac{u}{RC} \equiv f(x, u(x))$$

Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x,y), \; y(x_0) = y_0$

$$\frac{du}{dx} = -\frac{u}{RC} \quad ; \quad u(x_0) = u_0$$



Mètode d'Euler explícit o progressiu

1. Dividim l'interval $[x_0, x_f]$ en n subintervals equiespaiats $\{x_i\}_{i=0}^{n}$, amb $h = \dfrac{x_f - x_0}{n}$:

$$x_i = x_0 + ih, \; 0 \le i \le n, \qquad (16)$$

Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x, y),\ y(x_0) = y_0$

$$\frac{du}{dx} = -\frac{u}{RC} \quad ; \quad u(x_0) = u_0$$

$$\Updownarrow$$

Diferència finita progressiva:

$$\frac{u(x_1) - u(x_0)}{x_1 - x_0} = -\frac{u(x)}{RC}$$

On avaluem $f(x, u(x))$?

Per a obtenir un mètode *explícit* ho fem on coneixem les dades, és a dir, a x_0:

$$\frac{u(x_1) - u(x_0)}{x_1 - x_0} = -\frac{u(x_0)}{RC}$$

o equivalentment:

$$u(x_1) = u(x_0) + h\frac{-u(x_0)}{RC}$$

Mètode d'Euler explícit o progressiu

1. Dividim l'interval $[x_0, x_f]$ en n subintervals equiespaiats $\{x_i\}_{i=0}^{n}$, amb $h = \dfrac{x_f - x_0}{n}$:
$$x_i = x_0 + ih,\ 0 \le i \le n, \tag{16}$$
2. Diferenciació numèrica: diferència finita progressiva de primer ordre Eq.(6).

Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x,y),\ y(x_0) = y_0$

$$\frac{du}{dx} = -\frac{u}{RC} \qquad ; \qquad u(x_0) = u_0$$

$$\begin{aligned}
u(x_1) &= u(x_0) + h\frac{-u(x_0)}{RC}\\
u(x_2) &= u(x_1) + h\frac{-u(x_1)}{RC}\\
u(x_3) &= u(x_2) + h\frac{-u(x_2)}{RC}\\
&\vdots\\
u(x_{i+1}) &= u(x_i) + h\frac{-u(x_i)}{RC}\\
&\vdots\\
u(x_n) &= u(x_{n-1}) + h\frac{-u(x_{n-1})}{RC}
\end{aligned}$$

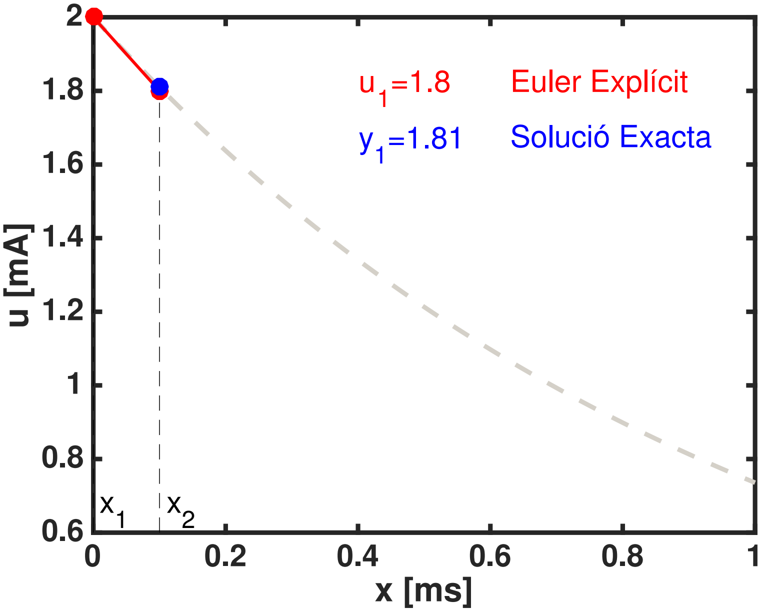
Utilitzarem indistintament $u(x_i)$ o u_i per a referir-nos a la solució numèrica en x_i.

Mètode d'Euler explícit o progressiu

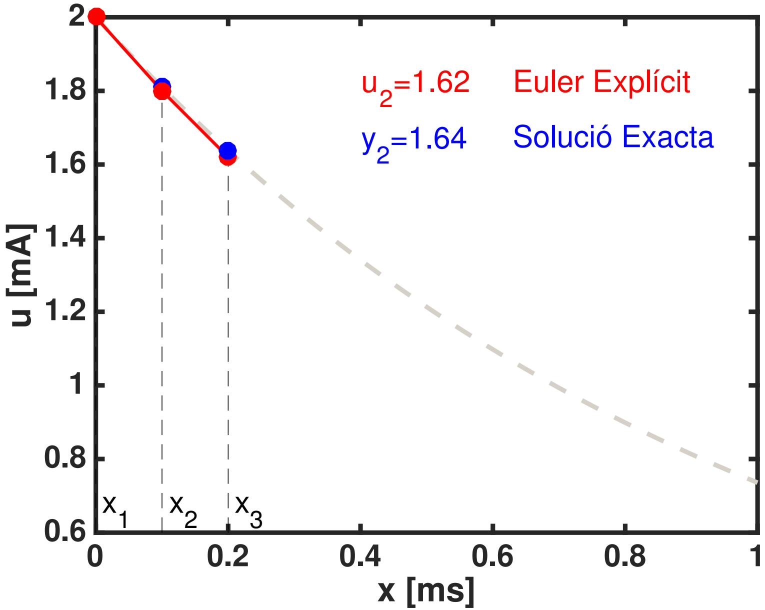
1. Dividim l'interval $[x_0, x_f]$ en n subintervals equiespaiats $\{x_i\}_{i=0}^{n}$, amb $h = \dfrac{x_f - x_0}{n}$:
$$x_i = x_0 + ih,\ 0 \le i \le n, \qquad (16)$$
2. Diferenciació numèrica: diferència finita progressiva de primer ordre Eq.(6).
3. Repetint el procediment anterior, es genera la successió d'aproximacions:

$$\begin{aligned}
u_1 &= u_0 + hf(x_0, u_0)\\
u_2 &= u_1 + hf(x_1, u_1)\\
&\vdots\\
u_{i+1} &= u_i + hf(x_i, u_i)\\
&\vdots\\
u_n &= u_{n-1} + hf(x_{n-1}, u_{n-1})
\end{aligned}$$

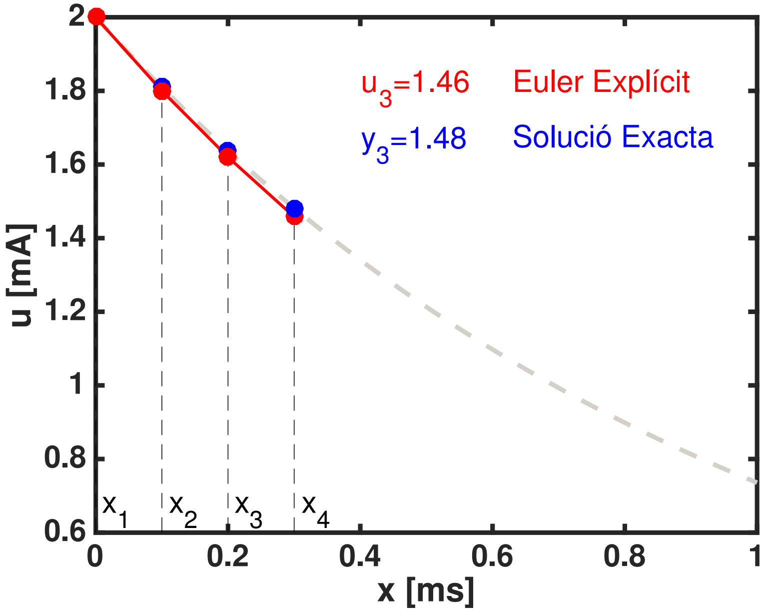
Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x, y),\ y(x_0) = y_0$



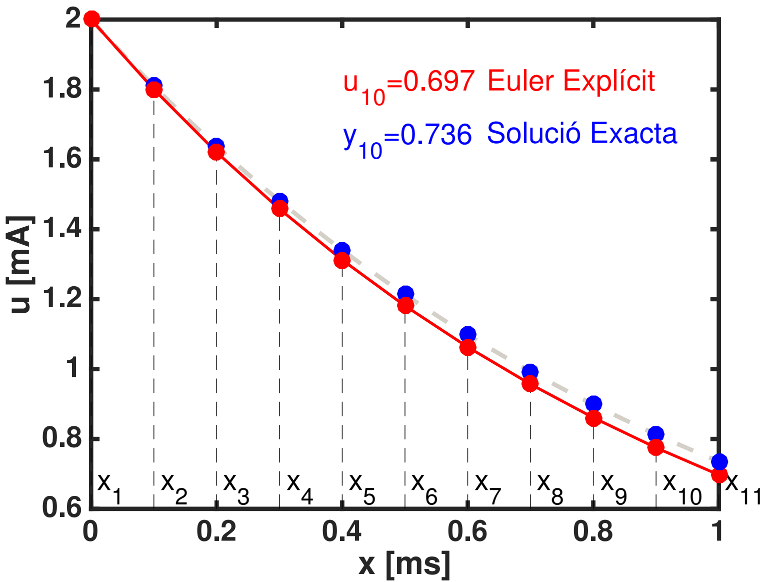
Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x, y),\ y(x_0) = y_0$



Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x, y),\ y(x_0) = y_0$



Mètode d'Euler (I): Algorisme de resolució del PVI: $y'(x) = f(x, y),\ y(x_0) = y_0$



Mètode d'Euler (III): Algorisme implícit o regressiu

En el mètode d'Euler explícit hem substituït la derivada $y'(x)$ per una diferència finita progressiva de primer ordre [vegeu Eq.(6)].

Anàlogament podem substituir $y'(x)$ per una diferència finita regressiva de primer ordre [vegeu Eq.(10)], la qual cosa dóna lloc al *mètode d'Euler implícit o regressiu*.

Mètode d'Euler implícit o regressiu (*backward Euler*)

1. Es genera la successió d'aproximacions següent:

$$
\begin{aligned}
u_1 &= u_0 + hf(x_1, u_1)\\
u_2 &= u_1 + hf(x_2, u_2)\\
&\vdots\\
u_{i+1} &= y_{i+1} + hf(x_{i+1}, u_{i+1})\\
&\vdots\\
y_n &= y_{n-1} + hf(x_n, y_n)
\end{aligned}
$$

2. Noteu que ara la funció $f(x_i, u_i)$ és avaluada *a partir de dades desconegudes*. En altres paraules, qualsevol dels u_i de la successió anterior s'obté a partir d'una relació implícita.

   Exemple: $u_2 = u_1 + hf(x_2, u_2)$

   Per tant, si $f(x_i, u_i)$ és una funció no lineal en u_i, en cada pas del mètode haurem de resoldre una equació no lineal en u_i.

   ⇒ És per aquesta raó que els mètodes implícits són computacionalment més costosos.

   ⇒ Els mètodes implícits tenen millors propietats d'estabilitat de la solució que els explícits, i permeten l'ús de valors de h grans.

Mètode d'Euler (IV): Exemple (I)

Siga el següent problema de valors inicials:

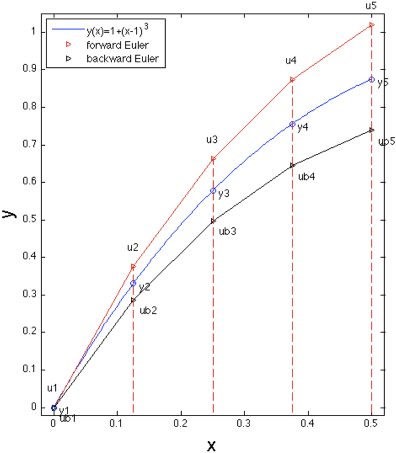
$$\text{PVI}: \begin{cases} y'(x) = 3(x-1)^2 \\ y(0) = 0 \\ y(0.5) = ? \end{cases}$$

Solució:

1. Calculem h prenent $n = 4 \implies h = \dfrac{0.5 - 0}{4} = 0.125$

La condició inicial $y(x_0) = y_0 \equiv u_0 \implies P_0(x_0 = 0, u_0 = 0)$

2. A partir de les dades inicials anteriors avaluem la següent taula:



i	$y'(x)$	x_i	u_i
0		0	0
1	$3(0-1)^2 = 3.000$	$0.000 + 0.125 = 0.125$	$0.000 + 0.125 * 3.000 = 0.375$
2	$3(0.125-1)^2 = 2.297$	$0.125 + 0.125 = 0.250$	$0.375 + 0.125 * 2.297 = 0.662$
3	$3(0.250-1)^2 = 1.688$	$0.250 + 0.125 = 0.375$	$0.662 + 0.125 * 1.688 = 0.873$
4	$3(0.375-1)^2 = 1.172$	$0.375 + 0.125 = 0.500$	$0.873 + 0.125 * 1.172 = 1.020$

Exercici 2: Repetiu l'exercici anterior utilitzant el mètode implícit d'Euler.

Mètode d'Euler (V): Errors

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició

$u_n^* \equiv$ solució numèrica en t_n que obtindríem partint de la solució exacta en t_{n-1}:

$$u_n^* = y_{n-1} + hf(t_{n-1}, y_{n-1}) \qquad (17)$$

Notació

u_n : Solució numèrica en t_n.

y_n : Solució exacta en t_n.

Errors en resoldre numèricament un PVI

Error local de discretització (comès en 1 pas del mètode),

$$e_n = y_n - u_n \qquad (18)$$

Definició:

Error global de discretització (màxim comès en qualsevol dels passos del mètode)

$$e \equiv \max\{e_n\} \qquad (19)$$

Es pot demostrar que en el mètode d'Euler l'**error local de discretització** està **acotat**:

$$|e_n| \leq e^{L(t_n-t_0)}|e_0| + \underbrace{\frac{e^{L(t_n-t_0)}-1}{L}\frac{M}{2}}_{\text{constant}} h$$

$$\forall n = 0, \ldots, N$$

$$M \equiv \max_{t\in[t_0,t_N]} |y''(t)|$$

Mètode d'Euler (V): Errors

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició

$u_n^* \equiv$ solució numèrica en t_n que obtindríem partint de la solució exacta en t_{n-1}:

$$u_n^* = y_{n-1} + hf(t_{n-1}, y_{n-1}) \qquad (17)$$

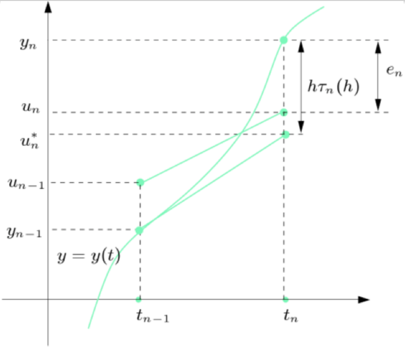
Errors en resoldre numèricament un PVI

Error local de discretització (comès en 1 pas del mètode),

$$e_n = y_n - u_n = (y_n - u_n^*) + (u_n^* - u_n), \quad (18)$$

es pot descompondre en:

1. $y_n - u_n^* \equiv$ error produït per un sol pas del mètode (**truncament**).



Mètode d'Euler (V): Errors

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició

$u_n^* \equiv$ solució numèrica en t_n que obtindríem partint de la solució exacta en t_{n-1}:

$$u_n^* = y_{n-1} + hf(t_{n-1}, y_{n-1}) \qquad (17)$$

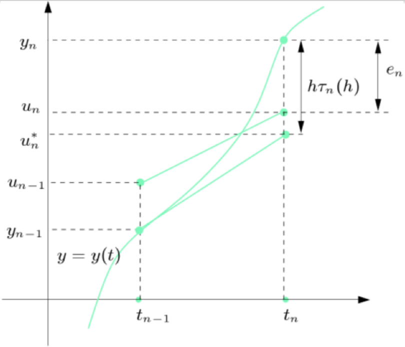
Errors en resoldre numèricament un PVI

Error local de discretització (comès en 1 pas del mètode),

$$e_n = y_n - u_n = (y_n - u_n^*) + (u_n^* - u_n), \quad (18)$$

es pot descompondre en:

1. $y_n - u_n^* \equiv$ error produït per un sol pas del mètode (**truncament**).
2. $u_n^* - u_n \equiv$ **propagació**, des de t_{n-1} fins a t_n, l'error acumulat en el pas de temps anterior t_{n-1}.



Mètode d'Euler (V): Errors

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició

$u_n^* \equiv$ solució numèrica en t_n que obtindríem partint de la solució exacta en t_{n-1}:

$$u_n^* = y_{n-1} + hf(t_{n-1}, y_{n-1}) \qquad (17)$$

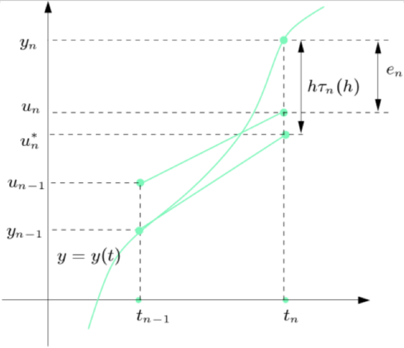
Errors en resoldre numèricament un PVI

Error local de discretització (comès en 1 pas del mètode),

$$e_n = y_n - u_n = (y_n - u_n^*) + (u_n^* - u_n), \quad (18)$$

es pot descompondre en:

1. $y_n - u_n^* \equiv$ error produït per un sol pas del mètode (**truncament**).
2. $u_n^* - u_n \equiv$ **propagació**, des de t_{n-1} fins a t_n, l'error acumulat en el pas de temps anterior t_{n-1}.
3. **arrodoniment**: deguts a la precisió (finita) dels càlculs i la representació dels valors en un SPF.



Mètode d'Euler (VI): Exemple (II)

Siga el següent problema de valors inicials:

$$PVI = \begin{cases} \frac{dy}{dx} = \sin x - \log y \\ y(0.13) = 0.32 \\ y(0.14) = ? \end{cases}$$

Solució:

1. Calculem el valor de h prenent $n = 4 \implies h = \frac{0.14-0.13}{4} = 0.0025$

La condició inicial $y(x_0) = y_0 \implies P_0 \ (x_0 = 0.13\,,\ y_0 = 0.32)$

2. A partir de les dades inicials anteriors:

$\frac{dy}{dx}$	x	y
	0.13	0.32
$\sin(0.13) - \log(0.32) = 1.269068$	$0.13 + 0.0025 = 0.1325$	$0.32 + 0.0025 * 1.269068 = 0.323172$
$\sin(0.1325) - \log(0.323172) = 1.261681$	$0.1325 + 0.0025 = 0.135$	$0.323172 + 0.0025 * 1.261681 = 0.326326$
$\sin(0.135) - \log(0.326326) = 1.254446$	$0.135 + 0.0025 = 0.1375$	$0.326326 + 0.0025 * 1.254446 = 0.329462$
$\sin(0.1375) - \log(0.329462) = 1.247358$	$0.1375 + 0.0025 = 0.14$	$0.329462 + 0.0025 * 1.247358 = 0.33258138$

$\implies$ Valor calculat: $y_4 = 0.33258138$

3. *Solució exacta*: $y(0.14) = 0.3325459 \implies \epsilon_r = \left\| \frac{0.33258138 - 0.3325459}{0.3325459} \right\| \approx 10^{-4} = 10^{-2}\,\%$

- Com que l'aproximació d'una corba $y(x)$ mitjançant una col·lecció de segments de recta no és exacta, es comet un error derivat del mètode
  $\implies$ **error de truncament**.

- Aquest error es pot disminuir reduint el valor de h, però s'obtindrà un nombre més alt de càlculs i, per tant, un error d'arrodoniment molt més gran.

Estabilitat (I)

Burden & Faires §5.10

Definició: Estabilitat

Un mètode diem que és estable si els seus resultats es basen *contínuament* en les dades inicials.

- En aplicacions pràctiques, tant les condicions inicials com les operacions aritmètiques que s'efectuen després són afectades per errors d'arrodoniment.
- Fins i tot si un mètode és convergent, és possible que petites pertorbacions en les condicions inicials ocasionen enormes diferències en les aproximacions posteriors (mètode inestable).
- El que ens interessa en la pràctica és que variacions petites de les condicions inicials produïsquen canvis igualment petits en la solució numèrica (mètode estable).

En el cas del mètode d'Euler podem escriure la condició d'estabilitat com un criteri senzill sobre la mida de h:

Condició d'estabilitat

$$h < \frac{2}{\max\limits_{t\in[t_0,t_n]} \left|\frac{\partial f}{\partial y}(t, y(t))\right|} \qquad (20)$$

Estabilitat (II): Exemple (Euler explícit)

Siga el següent problema de valors inicials:

$$\text{PVI}: \begin{cases} y'(t) = -ky \\ y(0) = 1 \\ k = 100 \end{cases}$$

Solució exacta: $y(t) = e^{-kt}$.

Apliquem el mètode d'Euler per a un pas h que després especificarem:

$$y_{n+1} = y_n + h(-ky_n) = (1 - hk)y_n$$

Aplicant de forma reiterada aquesta fórmula podem expressar la solució aproximada en funció de y_0:

$$y_{n+1} = (1 - hk)^{n+1} y_0, \quad n = 1, \ldots, N$$

La solució discreta (aproximada) serà *estable* si $|1 - hk| \leq 1$, és a dir, si $h \leq h_{\max} = 2/|k|$.[1] Per tant, si k és prou gran, podem trobar-nos que els valors de h que puguem prendre siguen minúsculs. Vegem alguns exemples:

$h = 1$			
i	$y(x)$	x_i	e^{-100t}
0	1	0	1
1	-99	1	3.72008×10^{-44}
2	9801	2	1.3839×10^{-87}
3	-970299	3	5.1482×10^{-131}
4	9.60596×10^{7}	4	1.91517×10^{-174}
5	-9.5099×10^{9}	5	7.12458×10^{-218}

$h = 0.05$			
i	$y(x)$	x_i	e^{-100t}
0	1	0	1
1	-4	0.05	0.00673795
10	1.04858×10^{6}	0.5	1.92875×10^{-22}
20	1.09951×10^{12}	1	3.72008×10^{-44}
30	1.15292×10^{18}	1.5	7.1751×10^{-66}
40	1.20893×10^{24}	2	1.3839×10^{-87}
50	1.26765×10^{30}	2.5	2.66919×10^{-109}
101	-6.42775×10^{60}	5.05	4.8005×10^{-220}

[1] Aquest valor de $h_{\max}$ coincideix amb el que s'obté de la condició [20], ja que: $\left|\frac{\partial f}{\partial y}(t, y(t))\right| = |k|$

Estabilitat (III): Exemple (Euler implícit)

Siga el següent problema de valors inicials:

$$\text{PVI}: \begin{cases} y'(t) = -ky \\ y(0) = 1 \\ k = 100 \end{cases}$$

Solució exacta: $y(t) = e^{-kt}$.

Repetim l'exercici anterior utilitzant el mètode implícit d'Euler. En aquest cas, tenim

$$y_{n+1} = y_n + h(-k y_{n+1}) \Longrightarrow y_{n+1} = \frac{y_n}{(1+hk)}$$

Aplicant de forma reiterada aquesta fórmula podem expressar la solució aproximada en funció de y_0:

$$y_{n+1} = \frac{y_0}{(1+hk)^{n+1}}, \quad n = 1, \ldots, N$$

La solució discreta (aproximada) serà *estable* independentment del valor del pas h (diem que és *incondicionalment estable*), encara que la precisió amb la qual calculem la solució decreix en augmentar h:

$h = 2$			
i	$y(x)$	x_i	e^{-100t}
0	1	0	1
1	4.975124×10^{-3}	2.0	1.383897×10^{-87}
2	2.475186×10^{-5}	4.0	$1.915170 \times 10^{-174}$
3	1.231436×10^{-7}	6.0	$2.650397 \times 10^{-261}$

$h = 1$			
i	$y(x)$	x_i	e^{-100t}
0	1	0	1
1	9.900990×10^{-3}	1.0	3.720076×10^{-44}
2	9.802960×10^{-5}	2.0	1.383897×10^{-87}
3	9.705901×10^{-7}	3.0	$5.148200 \times 10^{-131}$
4	9.609803×10^{-9}	4.0	$1.915170 \times 10^{-174}$
5	9.514657×10^{-11}	5.0	$7.124576 \times 10^{-218}$
6	9.420452×10^{-13}	6.0	$2.650397 \times 10^{-261}$

Estabilitat (IV): Exemple (III)

Exemple: Considerem el PVI;

$$\begin{cases} y'(t) = 1 - y(t)^2, & t > 0, \\ y(0) = (e-1)/(e+1), \end{cases} \tag{21}$$

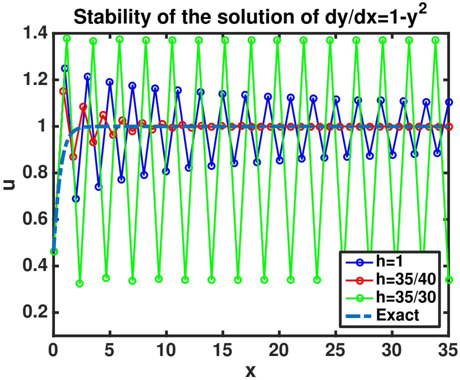
la solució analítica del qual és $y(t) = (e^{2t+1} - 1)/(e^{2t+1} + 1)$.

1. Quin seria el valor màxim de h perquè la resolució del PVI fóra estable?

$$\frac{\partial f}{\partial y}(t, y(t)) = -2y \quad \text{es pot demostrar que} \quad 0.9 \leq \left| \frac{\partial f}{\partial y}(t, y(t)) \right| \leq 2 \quad \text{si } t > 0$$

Hem de fer servir la condició d'estabilitat:

$$h < 2 \Big/ \max_{t \in [t_0, t_n]} \left| \frac{\partial f}{\partial y}(t, y(t)) \right| = 2/2 = 1 \equiv h_0$$



Estabilitat (V): Exemple (IV)

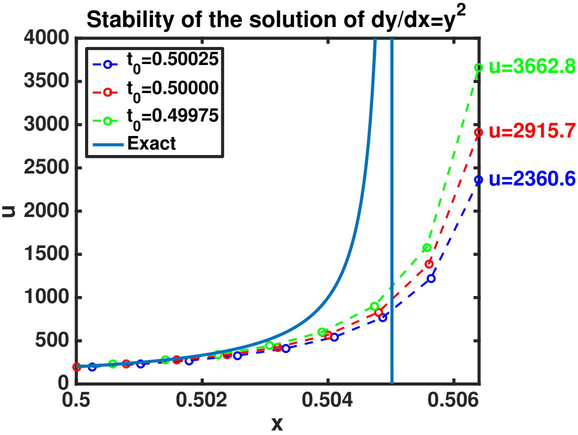
Exemple: Considereu el PVI;

$$\begin{cases} y'(t) = y(t)^2, & t \in [0.5, 0.5064], \\ y(0) = 200, \end{cases} \tag{22}$$

la solució analítica del qual és $y(t) = 0.5/(1 - 0.5(t - 0.5))$.

Com de sensible és la solució a les condicions inicials?

En la gràfica veiem el resultat de calcular la solució quan variem lleugerament (un $0.05\,\%$) al voltant de $t_0 = 0.5$. Es pot observar que després de tan sols $T = 0.5064$, les diferents solucions difereixen en més d'un $25\,\%$.



Consistència (I)

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició: Error local de truncament

$$\tau_n(h) \equiv \frac{y_n - u_n^*}{h}$$

En el cas del mètode d'Euler explícit: $|\tau_n(h)| \leq Mh/2$, e.d., l'ordre del mètode coincideix amb el del seu error local de truncament.

Important

Utilitzant (17):
$$\frac{y_n - y_{n-1}}{h} = f(t_{n-1}, y_{n-1}) + \tau_n(h)$$

L'error local de truncament representa allò que menyspreem en discretitzar la derivada.

També podem escriure:

$$y_n = y_{n-1} + hf(t_{n-1}, y_{n-1}) + h\tau_n(h)),$$

on es veu que $h\tau_n(h)$ és el que *li falta* a la solució exacta per a verificar el mètode d'Euler.

Consistència (II)

Definició: Error global de truncament o error global

$$\tau(h) = \max_{n=0,\dots,N} |\tau_n(h)| \tag{23}$$

En el cas del mètode d'Euler progressiu l'error global és

$$\tau(h) = \frac{Mh}{2}, \qquad M = \max_{t\in[t_0,t_N]} |y''(t)| \tag{24}$$

Si M és finit, és a dir, si $y''(t)$ està acotada a l'interval d'integració:

$\Longrightarrow$ L'error global del mètode progressiu d'Euler també està acotat.

$\Longrightarrow$ L'error global diminueix quan $h \to 0$.

Definició: Consistència

Un mètode es diu que és consistent si $\lim_{h\to 0} \tau(h) = 0$.

Direm que un mètode és consistent amb ordre p si $\tau(h) = \mathcal{O}(h^p)$, sent p un enter tal que $p \geq 1$.

Convergència (I)

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició: **Convergència**

Un mètode és convergent si

$$\forall n = 0, \ldots, N \quad |e_n| = |y_n - u_n| \leq C(h) \tag{25}$$

sent $C(h)$ infinitèsim repecte de h, quan $h \to 0$.

Si $C(h) = \mathcal{O}(h^p)$ per a algun $p \geq 1$ es diu que el mètode convergeix amb ordre p.

Convergència (I)

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició: Convergència

Un mètode és convergent si

$$\forall n = 0, \ldots, N \quad |e_n| = |y_n - u_n| \leq C(h) \tag{25}$$

sent $C(h)$ infinitèsim repecte de h, quan $h \to 0$.

Si $C(h) = \mathcal{O}(h^p)$ per a algun $p \geq 1$ es diu que el mètode convergeix amb ordre p.

Definició (2): Convergència

Un mètode és convergent si l'error global de discretització (19) tendeix a zero si $h \to 0$.

Convergència (I)

Quarteroni, et al., §7.2.1; Burden & Faires, def. 5.11

Definició: Convergència

Un mètode és convergent si

$$\forall n = 0, \ldots, N \quad |e_n| = |y_n - u_n| \leq C(h) \tag{25}$$

sent $C(h)$ infinitèsim repecte de h, quan $h \to 0$.

Si $C(h) = \mathcal{O}(h^p)$ per a algun $p \geq 1$ es diu que el mètode convergeix amb ordre p.

Definició (2): Convergència

Un mètode és convergent si l'error global de discretització (19) tendeix a zero si $h \to 0$.

Convergència del mètode d'Euler

Com que

$$|e_n| \leq \text{constant}_1 \times |e_0| + \text{constant}_2 \times h$$

el mètode d'Euler progressiu convergeix amb ordre 1 (si $|e_0| = 0$).

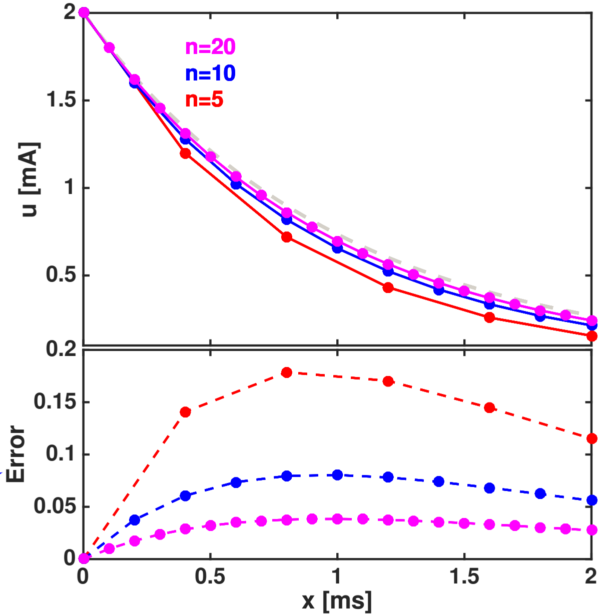
Convergència (II)

Considerem novament l'exemple del circuit, resolt amb el mètode d'Euler utilitzant diferents valors de n ($n = 5, 10, 20$).

Convergència: Incrementant n ($h \to 0$), la solució obtinguda amb el mètode d'Euler *s'assembla cada vegada més* (=*convergeix*) a la solució analítica.

Valor absolut de l'error:

$$|e_i| = |y_i - u_i|$$



Relació Convergència - Consistència - Estabilitat

CONSISTÈNCIA

$+$

ESTABILITAT

$\Longleftrightarrow$

CONVERGÈNCIA

Continguts

1 Introducció

2 Diferenciació numèrica

3 El problema de Cauchy o de valors inicials

4 Resolució numèrica de PVIs
- Mètode d'Euler
- Mètodes Runge-Kutta

Mètodes Runge-Kutta (I). *Problema de valors inicials (PVI):* $\frac{dy}{dx} = f(x,y) \; , \; y(x_0) = y_0$

¶¶ **Algorisme general:** $y_{i+1} = y_i + \phi(x_i, y_i, h)\ h \quad , \quad \phi = \sum_{j=1}^{n} a_j\ k_j$

$\phi(x_i, y_i, h)$ (funció increment) es pot interpretar com un **pendent representatiu** en l'interval (x_i, x_{i+1}), les a_i són constants $\left(\sum_{j=1}^{n} a_j = 1\right)$, i les k_j ($j-$**èsima estimació del pendent**) són definides per les relacions de recurrència:

$$
\begin{aligned}
k_1 &= f(x_i, y_i) \\
k_2 &= f(x_i + p_1\ h, y_i + q_{11}\ k_1\ h) \\
k_3 &= f(x_i + p_2\ h, y_i + q_{21}\ k_1\ h + q_{22}\ k_2\ h) \\
\dots &= \dots \\
k_n &= f(x_i + p_{n-1}\ h, y_i + \sum_{j=1}^{n-1} q_{n-1,j}\ k_j\ h)
\end{aligned}
$$

Mètodes Runge-Kutta (I). *Problema de valors inicials (PVI):* $\frac{dy}{dx} = f(x,y)$, $y(x_0) = y_0$

¶¶ **Algorisme general:** $y_{i+1} = y_i + \phi(x_i, y_i, h)\ h$, $\phi = \sum_{j=1}^{n} a_j\ k_j$

$\phi(x_i, y_i, h)$ (funció increment) es pot interpretar com un **pendent representatiu** en l'interval (x_i, x_{i+1}), les a_i són constants $(\sum_{j=1}^{n} a_j = 1)$, i les k_j ($j-$**èsima estimació del pendent**) són definides per les relacions de recurrència:

$$
\begin{aligned}
k_1 &= f(x_i, y_i) \\
k_2 &= f(x_i + p_1\ h, y_i + q_{11}\ k_1\ h) \\
k_3 &= f(x_i + p_2\ h, y_i + q_{21}\ k_1\ h + q_{22}\ k_2\ h) \\
\ldots &= \ldots \\
k_n &= f(x_i + p_{n-1}\ h, y_i + \sum_{j=1}^{n-1} q_{n-1,j}\ k_j\ h)
\end{aligned}
$$

¶ $\mathbf{n = 1}$ $\Longrightarrow$ *Mètode d'Euler.*

¶ Mètodes de Runge-Kutta (RK): Família de mètodes iteratius (implícits, explícits) per EDOs. **Aconsegueixen l'exactitud d'una sèrie de Taylor sense requerir el càlcul de derivades superiors.**

Mètodes Runge-Kutta (II): Segon ordre (I) Burden & Faires, §5.4.

¶¶ **Algorisme:**

$$y_{i+1} = y_i + \phi(x_i, y_i, h)h, \quad \phi = a_1 k_1 + a_2 k_2, \quad k_1 = f(x_i, y_i), \quad k_2 = f(x_i + p_1\, h, y_i + q_{11}\, k_1\, h) \tag{26}$$

¶¶ Els valors de a_1, a_2, p_1 i q_{11} són avaluats en igualar y_{i+1} amb el desenvolupament de Taylor (fins a ordre 2) $\Longrightarrow$ 3 equacions i 4 constants desconegudes:

$$a_1 \;+\; a_2 = 1 \quad , \quad a_2\; p_1 = 1/2 \quad , \quad a_2\; q_{11} = 1/2 \tag{27}$$

Suposem coneguda una de les incògnites, p. ex., a_2 $\Longrightarrow$ $\boxed{a_1 = 1 - a_2 \quad , \quad p_1 = q_{11} = 1/(2\, a_2)}$

¶ Com que es pot triar un nombre infinit de valors per a a_2 $\Longrightarrow$ Hi ha un nombre infinit de mètodes RK de segon ordre. Cadascun d'aquests mètodes donaria exactament els mateixos resultats si la solució de l'EDO fóra quadràtica, lineal o una constant.

Mètodes Runge-Kutta (II): Segon ordre (I) Burden & Faires, §5.4.

¶¶ **Algorisme:**

$$y_{i+1} = y_i + \phi(x_i, y_i, h)h, \quad \phi = a_1 k_1 + a_2 k_2, \quad k_1 = f(x_i, y_i), \quad k_2 = f(x_i + p_1\, h, y_i + q_{11}\, k_1\, h) \quad (26)$$

¶¶ Els valors de a_1, a_2, p_1 i q_{11} són avaluats en igualar y_{i+1} amb el desenvolupament de Taylor (fins a ordre 2) $\Longrightarrow$ 3 equacions i 4 constants desconegudes:

$$a_1 + a_2 = 1 \quad , \quad a_2\, p_1 = 1/2 \quad , \quad a_2\, q_{11} = 1/2 \quad (27)$$

Suposem coneguda una de les incògnites, p. ex., a_2 $\Longrightarrow$ $a_1 = 1 - a_2 \quad , \quad p_1 = q_{11} = 1/(2\, a_2)$

¶ Com que es pot triar un nombre infinit de valors per a a_2 $\Longrightarrow$ Hi ha un nombre infinit de mètodes RK de segon ordre. Cadascun d'aquests mètodes donaria exactament els mateixos resultats si la solució de l'EDO fóra quadràtica, lineal o una constant.

¶¶ **Casos particulars:**

$a_2 = 1/2$ $\Longrightarrow$ Mètode d'Heun amb un sol corrector

$$y_{i+1} = y_i + \left(\frac{k_1}{2} + \frac{k_2}{2}\right) h \quad , \quad k_1 = f(x_i, y_i) \quad , \quad k_2 = f(x_i + h, y_i + k_1\, h) \quad (28)$$

Mètodes Runge-Kutta (II): Segon ordre (I) Burden & Faires, §5.4.

¶¶ **Algorisme:**

$$y_{i+1} = y_i + \phi(x_i, y_i, h)h, \quad \phi = a_1 k_1 + a_2 k_2, \quad k_1 = f(x_i, y_i), \quad k_2 = f(x_i + p_1 h, y_i + q_{11} k_1 h) \quad (26)$$

¶¶ Els valors de a_1, a_2, p_1 i q_{11} són avaluats en igualar y_{i+1} amb el desenvolupament de Taylor (fins a ordre 2) $\Longrightarrow$ 3 equacions i 4 constants desconegudes:

$$a_1 + a_2 = 1 \quad , \quad a_2\, p_1 = 1/2 \quad , \quad a_2\, q_{11} = 1/2 \quad (27)$$

Suposem coneguda una de les incògnites, p. ex., a_2 $\Longrightarrow$ $a_1 = 1 - a_2 \quad , \quad p_1 = q_{11} = 1/(2\, a_2)$

¶ Com que es pot triar un nombre infinit de valors per a a_2 $\Longrightarrow$ Hi ha un nombre infinit de mètodes RK de segon ordre. Cadascun d'aquests mètodes donaria exactament els mateixos resultats si la solució de l'EDO fóra quadràtica, lineal o una constant.

¶¶ **Casos particulars:**

$a_2 = 1/2$ $\Longrightarrow$ Mètode d'Heun amb un sol corrector

$$y_{i+1} = y_i + \left(\frac{k_1}{2} + \frac{k_2}{2}\right) h \quad , \quad k_1 = f(x_i, y_i) \quad , \quad k_2 = f(x_i + h, y_i + k_1 h) \quad (28)$$

$a_2 = 1$ $\Longrightarrow$ Mètode del punt mitjà (també conegut com a mètode d'Euler millorat)

$$y_{i+1} = y_i + k_2 h \quad , \quad k_1 = f(x_i, y_i) \quad , \quad k_2 = f\left(x_i + \frac{h}{2}, y_i + k_1 \frac{h}{2}\right) \quad (29)$$

Mètodes Runge-Kutta (II): Segon ordre (I) Burden & Faires, §5.4.

¶¶ **Algorisme:**

$$y_{i+1} = y_i + \phi(x_i, y_i, h)h, \quad \phi = a_1 k_1 + a_2 k_2, \quad k_1 = f(x_i, y_i), \quad k_2 = f(x_i + p_1\, h, y_i + q_{11}\, k_1\, h) \tag{26}$$

¶¶ Els valors de a_1, a_2, p_1 i q_{11} són avaluats en igualar y_{i+1} amb el desenvolupament de Taylor (fins a ordre 2) $\Longrightarrow$ 3 equacions i 4 constants desconegudes:

$$a_1 + a_2 = 1 \quad , \quad a_2\, p_1 = 1/2 \quad , \quad a_2\, q_{11} = 1/2 \tag{27}$$

Suposem coneguda una de les incògnites, p. ex., $a_2 \Longrightarrow$ $\boxed{a_1 = 1 - a_2 \quad , \quad p_1 = q_{11} = 1/(2\, a_2)}$

¶ Com que es pot triar un nombre infinit de valors per a $a_2 \Longrightarrow$ Hi ha un nombre infinit de mètodes RK de segon ordre. Cadascun d'aquests mètodes donaria exactament els mateixos resultats si la solució de l'EDO fóra quadràtica, lineal o una constant.

¶¶ **Casos particulars:**

$a_2 = 1/2 \Longrightarrow$ Mètode d'Heun amb un sol corrector

$$y_{i+1} = y_i + \left(\frac{k_1}{2} + \frac{k_2}{2}\right) h \quad , \quad k_1 = f(x_i, y_i) \quad , \quad k_2 = f(x_i + h, y_i + k_1\, h) \tag{28}$$

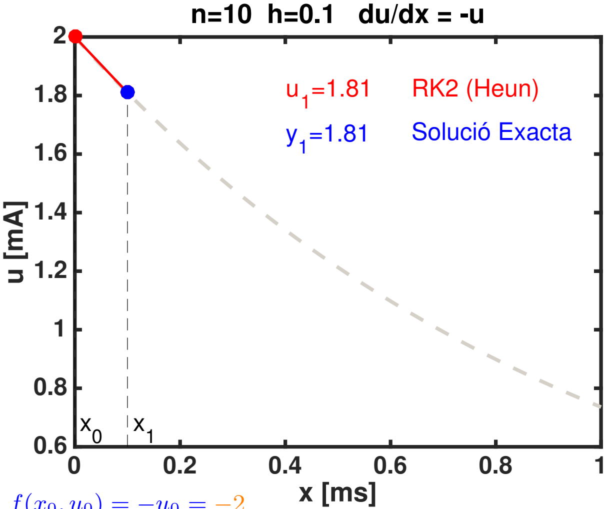
$a_2 = 1 \Longrightarrow$ Mètode del punt mitjà (també conegut com a mètode d'Euler millorat)

$$y_{i+1} = y_i + k_2\, h \quad , \quad k_1 = f(x_i, y_i) \quad , \quad k_2 = f\left(x_i + \frac{h}{2}, y_i + k_1 \frac{h}{2}\right) \tag{29}$$

$a_2 = 2/3 \Longrightarrow$ Mètode de Ralston

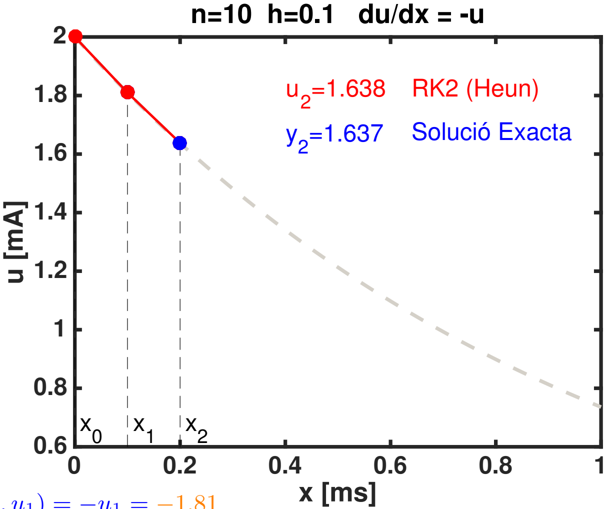
$$y_{i+1} = y_i + \left(\frac{k_1}{3} + \frac{2k_2}{3}\right) h \quad , \quad k_1 = f(x_i, y_i) \quad , \quad k_2 = f\left(x_i + \frac{3h}{4}, y_i + 3k_1 \frac{h}{4}\right) \tag{30}$$

Mètodes Runge-Kutta (II): RK2 Exemple



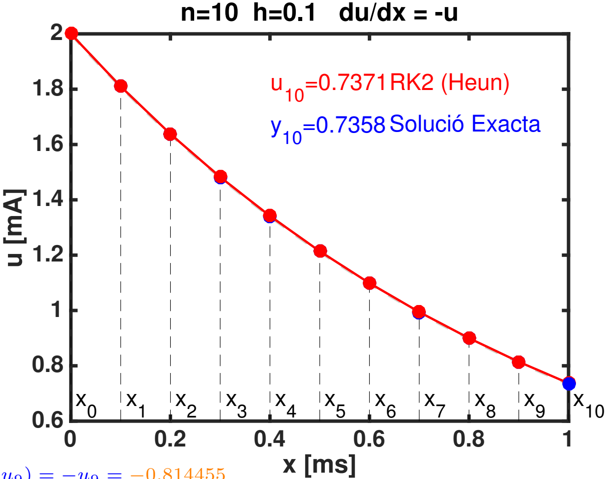
$$
\begin{aligned}
k_1 &= f(x_0, u_0) = -u_0 = -2 \\
k_2 &= f(x_0 + h, u_0 + k_1 h) = f(x_1, 2 + (-2) \times 0.1) = f(0.1, 1.8) = -1.8 \\
u_1 &= u_0 + \frac{k_1 + k_2}{2} h = 2 + \frac{-2 + (-1.8)}{2} \times 0.1 = 1.81;
\end{aligned}
$$

Mètodes Runge-Kutta (II): RK2 Exemple



$$
\begin{aligned}
k_1 &= f(x_1, u_1) = -u_1 = -1.81 \\
k_2 &= f(x_1 + h, u_1 + k_1 h) = f(x_2, 1.81 + (-1.81) \times 0.1) = f(0.2, 1, 629) = -1.629 \\
u_2 &= u_1 + \frac{k_1 + k_2}{2} h = 1.81 + \frac{-1.81 + (-1.629)}{2} \times 0.1 = 1.63805;
\end{aligned}
$$

Mètodes Runge-Kutta (II): RK2 Exemple



$$
\begin{aligned}
k_1 &= f(x_9, u_9) = -u_9 = -0.814455 \\
k_2 &= f(x_9 + h, u_9 + k_1 h) = f(x_{10}, 0.814455 + (-0.814455) \times 0.1) = f(0.9, 0.73301) = -0.73301 \\
u_{10} &= u_9 + \frac{k_1 + k_2}{2} h = 0.814455 + \frac{-0.814455 + (-0.73301)}{2} \times 0.1 = 0.737082;
\end{aligned}
$$

Mètodes Runge-Kutta (VI): Quart ordre

¶¶ **Algorisme:** Per a n = 4, estenent el raonament usat en els casos anteriors (n = 2, 3) $\Longrightarrow$ Un membre de la família dels mètodes RK denominat **RK4** és:

Algorisme RK4

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h$$

on

$$\begin{aligned}
k_1 &= f\left(x_i, y_i\right) \\
k_2 &= f\left(x_i + \tfrac{1}{2}h, y_i + \tfrac{1}{2}k_1 h\right) \\
k_3 &= f\left(x_i + \tfrac{1}{2}h, y_i + \tfrac{1}{2}k_2 h\right) \\
k_4 &= f\left(x_i + h, y_i + k_3 h\right)
\end{aligned}$$

Mètodes Runge-Kutta (VII): Exemple (I)

Siga el PVI: $\frac{dy}{dx} = f(x,y)$, $f(x,y) = \frac{1}{2}(1+x)\ y^2$, $y(x_0 = 0) = 1$, $I \in [0, 0.5]$

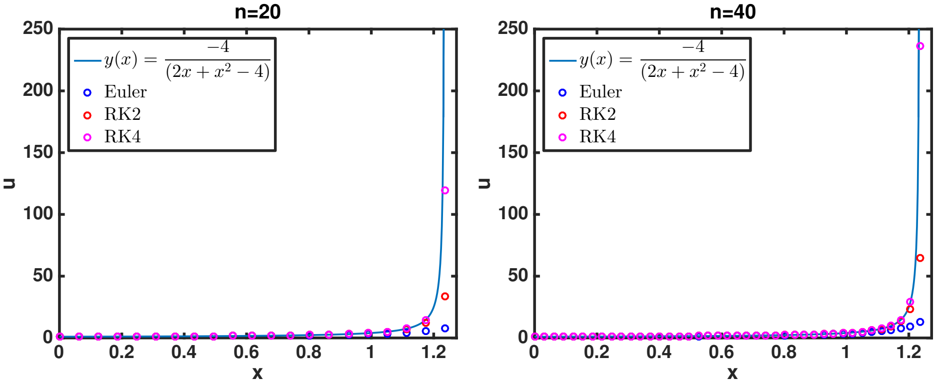
Solució: Prenguem $h = 0.1$. Sabem que la solució exacta és $y_{ex}(x) = -4/(2x + x^2 - 4)$.

x_i	y_i	$k_1 h$	$k_2 h$	$k_3 h$	$k_4 h$	y_{ex}
0	1	0.05	0.05515781	0.05543572	0.06126695	1
0.1	1.055409	0.06126385	0.03538455	0.06621391	0.07548228	1.05540897
0.2	1.11206618	0.07420147	0.07923508	0.08289846	0.09281613	1.12359551
0.3	1.19394696	0.09265811	0.0999885	0.10444881	0.11800821	1.20845921
0.4	1.29720378	0.11779163	0.12873043	0.13440558	0.1537129	1.31578947
0.5	1.43016654	0.15340322	0.17029887	0.17795414	0.20688417	1.45454545

Mètodes Runge-Kutta (VII): Exemple (I)

PVI: $\dfrac{dy}{dx} = f(x,y)$, $f(x,y) = \dfrac{1}{2}(1+x)\,y^2$, $y(x_0=0)=1$, $I \in [0, 1.236]$

Hem estès l'interval on avaluem la solució i comparem els tres mètodes anteriors:



Si la solució és suau, tots els mètodes donen resultats *semblants*, però quan ens acostem a regions on la funció solució ($y(x)$) canvia ràpidament (p.e. en $x = -1 + \sqrt{5} \simeq 1.236$), podem apreciar amb claredat la diferència de precisió de cada mètode.

Matemàtiques III:
6. Inferència estadística

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Inferència i decisió

¶¶ **Estadística**: Ciència, amb base matemàtica, que tracta sobre la recol·lecció, anàlisi i interpretació de dades, i que busca explicar i/o deduir condicions regulars en fenòmens de tipus aleatori.

- *Estadística descriptiva*: Tècniques per a presentar i resumir un conjunt de dades, per a facilitar-ne la comprensió i/o per a generar hipòtesis.
- *Teoria de la probabilitat*: Branca de les matemàtiques que va nàixer per a explicar els jocs d'atzar i que constitueix la base de la inferència estadística.

¶¶ **Inferència estadística**: Comprèn dos elements bàsics:

- Tècniques que permeten deduir alguna propietat d'una població a partir de les dades d'una mostra extreta d'aquesta població.
- Processos de decisió sobre una població, basats en l'anàlisi estadística de dades procedents d'una o diverses mostres.

Exemple: Estimar l'alçada dels xiquets espanyols de 8 anys a partir d'una mostra de 5000 xiquets. Quina fiabilitat té l'estimació?

Inferència i Decisió: Conceptes bàsics (I)

¶¶ **Conceptes bàsics (I)** :

- **Població** Conjunt d'*individus* que es pretén estudiar.
  En l'exemple: La població és el conjunt dels xiquets espanyols de 8 anys.

- **Mostra** Subconjunt de la població del qual es tenen 'dades'
  En l'exemple: La mostra són els 5000 xiquets dels quals es disposa de la mesura de la seua alçada (les dades).

- **Variable** Una característica de la població que pren un valor per a cada individu d'aquesta. Les variables poden ser:
  1. categòriques (sexe, estat civil...)
  2. numèriques (temperatura, pes, alçada...)

  En l'exemple: La variable és l'alçada d'un xiquet de 8 anys.

- **Dades**: Valors que pot prendre una variable.
  En l'exemple: L'alçada pot prendre qualsevol valor positiu, tot i que sembla evident que valors *grans* seran molt poc probables.

Inferència i Decisió: Conceptes bàsics (II)

¶¶ Conceptes bàsics (II)

- **Experiment determinista**: Repetit en igualtat de condicions condueix al mateix resultat (situació ideal que no es produeix pràcticament mai)

  $\Rightarrow$ Les mateixes causes, en les mateixes condicions, produeixen els mateixos efectes.

  Exemple: Mesurar la intensitat de corrent en un circuit.

- **Experiment aleatori**: Repetit en igualtat de condicions no es pot predir el resultat

  $\Rightarrow$ Les *mateixes* causes, en les *mateixes* condicions, poden produir efectes diferents.

  Exemple: Llançar una moneda a l'aire.

¶ En un *experiment aleatori*, coneixem per endavant tots els seus possibles resultats, però no podem predir el resultat concret de l'experiment.

$\Longrightarrow$ El resultat d'un experiment aleatori és

- impredictible 'a priori', però
- pertany a un conjunt que es pot descriure completament abans de realitzar l'experiment.

Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Definicions

¶ El concepte de *variable aleatòria* sorgeix de la necessitat de fer més manejables matemàticament els resultats dels experiments aleatoris, que en molts casos són qualitatius i segueixen patrons molt similars, encara que la naturalesa de l'experiment no ho siga.

¶¶ Definició 1.- Siga Ω l'**espai mostral** associat a un experiment aleatori, és a dir, el conjunt de tots els possibles resultats d'aquest experiment. Una **variable aleatòria** és una aplicació: $X : \Omega \to \mathbb{R}$

¶¶ Definició 2.- Donada una variable aleatòria X, anomenarem **suport de X,** S_X, o **imatge de X,** Im_X, el conjunt de possibles valors de X a $\mathbb{R}$.

Observació: El suport d'una variable aleatòria pot ser discret (finit o numerable), o continu (un interval de $\mathbb{R}$, o tot $\mathbb{R}$)

¶¶ **Variables aleatòries discretes** $\Longrightarrow$ El seu suport és discret: $S_X = \{x_1, x_2, \ldots, x_n, \ldots\}$
¶¶ **Variables aleatòries contínues** $\Longrightarrow$ El seu suport és un interval de $\mathbb{R}$.

Exemples:

- Experiment 1: Llançar una moneda defectuosa amb $p(\text{cara}) = 0.9$
  $\Omega = \{\texttt{Cara}, \texttt{Creu}\} \iff S_X = \{0, 1\}$ (o qualssevol altres dos valors, p. ex., $S_X = \{2, 9\}$).
- Experiment 2: Observar una peça fabricada en un procés que produeix un 1 % de peces defectuoses.
  $\Omega = \{\texttt{bona}, \texttt{defectuosa}\} \iff S_X = \{0, 1\}$
- Experiment 3: Alçada d'una persona extreta d'una determinada població.
  $\Omega = [0, 270]\ \texttt{cm} \iff S_X = [0, 1]$ (o qualsevol altre interval, p.e., $S_X = [0, 270]$).

¶ El maneig de variables aleatòries permet 'classificar' els experiments aleatoris, de manera que el càlcul de probabilitats siga molt més senzill.

Continguts

1. Introducció
2. Variables aleatòries discretes: Definicions
3. Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat
4. Variables aleatòries discretes: Esperança i variància
5. Variables aleatòries discretes: Funcions densitat de probabilitat
6. Variables aleatòries contínues: Funcions densitat de probabilitat
7. Variables aleatòries contínues: Distribució de Gauss
8. Inferència

Llei de probabilitat per a variables aleatòries discretes (I)

¶¶ Una variable aleatòria (associada a un determinat experiment aleatori) trasllada la informació probabilística rellevant des de l'espai mostral Ω a $\mathbb{R}$, mitjançant una **probabilitat induïda** que es defineix sobre els conjunts de la recta real $\mathbb{R}$ de la següent manera: $P_X(A) = p(X^{-1}(A))$, amb $A \subset \mathbb{R}$

Exemple: Considerem l'experiment de llançar dos daus.

Variable aleatòria $X =$ suma de les cares

$$S_X = \{2, 3, 4, \dots, 12\} \qquad P_X(0) = p(X^{-1}(0)) = p(\emptyset) = 0$$

Variable aleatòria $Y =$ valor absolut de la diferència de les cares

$$S_Y = \{0, 1, \dots, 5\} \qquad P_Y(0) = p(Y^{-1}(0)) = \frac{6}{36} = \frac{1}{6}$$

ja que: $Y^{-1}(0) = \{\{1,1\}, \{2,2\}, \{3,3\}, \{4,4\}, \{5,5\}, \{6,6\}\}$

Llei de probabilitat per a variables aleatòries discretes (I)

¶¶ Una variable aleatòria (associada a un determinat experiment aleatori) trasllada la informació probabilística rellevant des de l'espai mostral Ω a $\mathbb{R}$, mitjançant una **probabilitat induïda** que es defineix sobre els conjunts de la recta real $\mathbb{R}$ de la següent manera: $P_X(A) = p(X^{-1}(A))$, amb $A \subset \mathbb{R}$

Exemple: Considerem l'experiment de llançar dos daus.

Variable aleatòria $X =$ suma de les cares

$$S_X = \{2, 3, 4, \dots, 12\} \qquad P_X(0) = p(X^{-1}(0)) = p(\emptyset) = 0$$

Variable aleatòria $Y =$ valor absolut de la diferència de les cares

$$S_Y = \{0, 1, \dots, 5\} \qquad P_Y(0) = p(Y^{-1}(0)) = \frac{6}{36} = \frac{1}{6}$$

ja que: $Y^{-1}(0) = \{\{1,1\}, \{2,2\}, \{3,3\}, \{4,4\}, \{5,5\}, \{6,6\}\}$

Un altre exemple: $P_Y(1)$?

$P_Y(1) = p(Y^{-1}(1)) = \frac{10}{36}$

$$\Omega = \left\{ \begin{array}{cccccccccc} \{1,1\} & \{2,2\} & \{3,3\} & \{4,4\} & \{5,5\} & \{6,6\} & & & & \\ \{1,2\} & \{2,1\} & \{2,3\} & \{3,2\} & \{3,4\} & \{4,3\} & \{4,5\} & \{5,4\} & \{5,6\} & \{6,5\} \\ \{1,3\} & \{3,1\} & \{2,4\} & \{4,2\} & \{3,5\} & \{5,3\} & \{4,6\} & \{6,4\} & & \\ \{1,4\} & \{4,1\} & \{2,5\} & \{5,2\} & \{3,6\} & \{6,3\} & & & & \\ \{1,5\} & \{5,1\} & \{2,6\} & \{6,2\} & & & & & & \\ \{1,6\} & \{6,1\} & & & & & & & & \end{array} \right\}$$

$Y^{-1}(1) = \{\{1,2\}, \{2,1\}, \{2,3\}, \{3,2\}, \{3,4\}, \{4,3\}, \{4,5\}, \{5,4\}, \{5,6\}, \{6,5\}\}$

Llei de probabilitat per a variables aleatòries discretes (I)

¶¶ Una variable aleatòria (associada a un determinat experiment aleatori) trasllada la informació probabilística rellevant des de l'espai mostral Ω a $\mathbb{R}$, mitjançant una **probabilitat induïda** que es defineix sobre els conjunts de la recta real $\mathbb{R}$ de la següent manera: $P_X(A) = p(X^{-1}(A))$, amb $A \subset \mathbb{R}$

Exemple: Considerem l'experiment de llançar dos daus.

Variable aleatòria $X =$ suma de les cares

$$S_X = \{2, 3, 4, \ldots, 12\} \qquad P_X(0) = p(X^{-1}(0)) = p(\emptyset) = 0$$

Variable aleatòria $Y =$ valor absolut de la diferència de les cares

$$S_Y = \{0, 1, \ldots, 5\} \qquad P_Y(0) = p(Y^{-1}(0)) = \frac{6}{36} = \frac{1}{6}$$

ja que: $Y^{-1}(0) = \{\{1,1\}, \{2,2\}, \{3,3\}, \{4,4\}, \{5,5\}, \{6,6\}\}$

¶¶ P_X és la **llei de probabilitat de X o distribució de probabilitat de X**.
La distribució de probabilitat de X, P_X, ens proporciona la informació necessària per a conèixer el comportament probabilístic de X.

¶¶ En la pràctica, resulta més convenient recórrer a la **funció de distribució de probabilitat acumulada (integrada)** d'una variable discreta (contínua), definida com a:

$$F_X : \mathbb{R} \quad \rightarrow \quad [0, 1]$$

$$F_X(x) = P_X((-\infty, x]) = P(X^{-1}\{(-\infty, x]\}) = P(X \leq x)$$

Llei de probabilitat per a variables aleatòries discretes (II)

Funcions distribució de probabilitat i densitat de probabilitat

¶¶ Per a una **variable aleatòria discreta**, la *llei de probabilitat* o *distribució de probabilitat* de X queda determinada pels valors $P_X(X = x_i)$, $i = 1, 2...$

¶¶ A partir d'aquests valors podem definir la **funció de densitat de probabilitat** de la variable aleatòria:

$$f_X(x) = \begin{cases} P_X(X = x_i), & x = x_i \\ 0, & x \neq x_i \end{cases}$$

La **funció de distribució** és aleshores:

$$F_X(x) = \sum_{x_i \leq x} P_X(X = x_i) = \sum_{x_i \leq x} f_X(x_i)$$

i per tant $F_X(x)$ és una funció escalonada, els salts es produeixen en els punts de S_X.

Llei de probabilitat per a variables aleatòries discretes (II): Exemple

Experiment aleatori: comptar el nombre de cares obtingudes en llançar 2 monedes:

- ¶¶ L'espai mostral està compost per quatre possibles resultats:

$$\Omega = \{\mathtt{cc}, \mathtt{cx}, \mathtt{xc}, \mathtt{xx}\}.$$

- ¶¶ El conjunt suport el formen els tres casos diferents que tenim (# cares):

$$S_X = \{0, 1, 2\}.$$

- ¶¶ La **funció de densitat de probabilitat** serà:

$$f_X(x) = \begin{cases} P_X(X=0) = 1/4, & \\ P_X(X=1) = 2/4, & \\ P_X(X=2) = 1/4, & \\ 0, & x \neq 0, 1, 2 \end{cases}$$

- ¶¶ La **funció de distribució** serà:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1/4, & 0 \leq x < 1 \\ 3/4, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Variables aleatòries discretes: Esperança

Esperança: definició

Esperança d'una variable aleatòria discreta, X, amb suport S_X:

$$E(X) = \sum_{x \in S_X} x P(X = x)$$

¶ En general, l'esperança és un **mitjana ponderada** dels valors que pren la variable, donant-li major pes als valors més probables. $E(X)$, en general, no coincideix amb un dels valors de S_X.

Exemple 1: X= nombre de cares en tirar dues vegades una moneda equilibrada.

$S_X = \{0, 1, 2\}$

x	0	1	2
$P(x)$	1/4	1/2	1/4

$E(X) = \sum_{i=0}^{2} iP(X = i) =$
$0 \cdot (1/4) + 1 \cdot (1/2) + 2 \cdot (1/4) = 1$

Exemple 2: X= nombre d'uns en tirar dues vegades un dau equilibrat.

$S_X = \{0, 1, 2\}$

x	0	1	2
$P(x)$	25/36	10/36	1/36

$E(X) = \sum_{i=0}^{2} iP(X = i) =$
$0 \cdot (25/36) + 1 \cdot (10/36) + 2 \cdot (1/36) = 1/3$

Exemple 3: Experiments aleatoris que consisteixen en l'extracció a l'atzar d'un individu en una mostra finita, p. ex., es tenen totes les notes de Matemàtiques III dels alumnes de primer.
$X =$ nota d'un alumne extret a l'atzar. $S_X = \{0, 1, \ldots, 10\}$
$P(X = x_i) = n_i/N$ amb $n_i =$ frequència de la nota x_i, $N =$ nombre total d'alumnes

$$E(X) = \sum_i x_i P(X = x_i) = \sum_i x_i \frac{n_i}{N} = \text{mitjana poblacional}$$

Variables aleatòries discretes: Variància

Variància: definició

Variància d'una variable aleatòria discreta, X, amb suport S_X:

$$Var(X) = \sum_{x \in S_X} (x - E(X))^2 P(X = x)$$

$\Longrightarrow$ La variància és una mesura de la dispersió.

Exemple 1: X= nombre de cares en tirar dues vegades una moneda equilibrada.

$S_X = \{0, 1, 2\}$

x	0	1	2
$P(x)$	1/4	1/2	1/4

$E(X) = 1$; $Var(X) = (0-1)^2 \frac{1}{4} + (1-1)^2 \frac{1}{2} + (2-1)^2 \frac{1}{4} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

Exemple 2: X= nombre d'uns en tirar dues vegades un dau equilibrat.

$S_X = \{0, 1, 2\}$

x	0	1	2
$P(x)$	25/36	10/36	1/36

$E(X) = 1/3$; $Var(X) = (0-1/3)^2 \frac{25}{36} + (1-1/3)^2 \frac{10}{36} + (2-1/3)^2 \frac{1}{36} = 0.2778$

Propietats bàsiques de l'esperança i de la variància Colomer, cap. 7

Notació $\sigma_X^2 \equiv Var(X)$

- Linealitat de l'esperança matemàtica. Si X i Y són dues variables aleatòries, i $a, b \in \mathbb{R}$, es compleix que

$$E(aX + bY) = aE(X) + bE(Y)$$

Demostració: L'espai mostral finit és $S = \{w_1, \ldots, w_n\}$. Anomenem $Z = aX + bY$. Aplicant la definició tenim:

$$\begin{aligned} E(Z) &= \sum_{j=1}^{n} [aX(w_j) + bY(w_j)]P(w_j) = \\ &= a\sum_{j=1}^{n} X(w_j)P(w_j) + b\sum_{j=1}^{n} Y(w_j)P(w_j) = aE(X) + bE(Y). \end{aligned}$$

- $\sigma_X^2 = E((X - E(X))^2)$.
- $\sigma_X^2 = E(X^2) - E(X)^2$.

Demostració:
$\sigma_X^2 = \sum_{j=1}^{n} (w_j - E(X))^2 P(w_j) = \sum_{j=1}^{n} w_j^2 P(w_j) - 2E(X)\sum_{j=1}^{n} w_j P(w_j) + E(X)^2 \sum_{j=1}^{n} P(w_j) = E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2$.

Propietats bàsiques de l'esperança i de la variància Colomer, cap. 7

Notació $\sigma_X^2 \equiv Var(X)$

- Siga X una variable aleatòria amb esperança $\mu = E(X)$ i variància σ_x^2. Siga la variable $Z = aX + b$, $a, b \in \mathbb{R}$. Es compleix que $\sigma_Z^2 = a^2\sigma_X^2$.

  Demostració:

  1. $E(Z) = E(aX + b) = aE(X) + b$
  2. $$\begin{aligned} \sigma_Z^2 &= E((Z - E(Z))^2) = E((aX + b - aE(X) - b)^2) = E((aX - aE(X))^2) = \\ &= a^2 E((X - E(X))^2) = a^2\sigma_X^2. \end{aligned}$$

- Si dues variables aleatòries X_1 i X_2 són independents, es compleix que:

$$\sigma_{X_1+X_2}^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2$$

$\Longrightarrow$ Corol·lari: Si n variables aleatòries $X_1, \ldots, X_n$ són independents:

$$\sigma_{X_1+X_2+\cdots+X_n}^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \cdots \sigma_{X_n}^2$$

Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

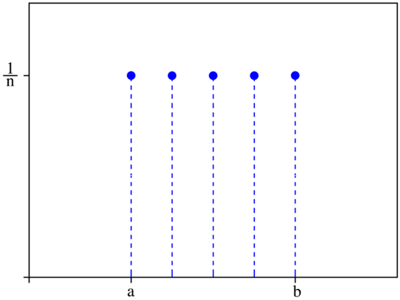
7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Distribució uniforme

Variable aleatòria discreta uniforme: $X \approx U(n)$

$$S_X = \{x_1, x_2, \ldots, x_n\}$$
$$P_X(X = x_i) = \frac{1}{n}$$



Experiment: Es realitza una prova que té un nombre finit de resultats equiprobables:

- monedes $n = 2$
- daus $n = 6$
- baralles $n = 40$
- ...

Distribució de Bernouilli

¶¶ Un experiment aleatori s'anomena de Bernouilli si verifica que:

- L'experiment consisteix a observar elements d'una mostra i classificar-los en dues categories: **Èxit (E)** i **Fracàs (F)**.
- La probabilitat d'èxit és p, i la de fracàs $q = 1 - p$ a cada realització de l'experiment.
- Les observacions són independents.

Variable aleatòria discreta de Bernouilli: $X \approx B(p)$

i) $S_X = \{0, 1\}$, $P_X(1) = p$, $P_X(0) = 1 - p$,

ii) La funció densitat de probabilitat d'aquesta distribució és (els valors discrets x els denotarem per $k \in \mathbb{N}$):

$$f_{B(p)}(x) := f(k;p) = \begin{cases} p & \text{si } k = 1, \\ 1 - p & \text{si } k = 0, \\ 0 & \text{si } k \neq 0, 1 \end{cases} \Longleftrightarrow f(k;p) = p^k (1-p)^{1-k} \quad \text{per a } k \in \{0, 1\}$$

iii) L'**esperança** d'una variable aleatòria de Bernoulli és $E(X) = p$

iv) La **variància** d'una variable aleatòria de Bernoulli és $\mathrm{Var}(X) = p(1 - p)$

Distribució binomial (I)

- L'experiment aleatori consisteix a repetir n vegades (n fix) 1 prova.
- En cada prova s'observa si surt o no un succés **E**.
- La probabilitat d'èxit val p i és la mateixa en totes les proves.
- El resultat d'una prova no influeix sobre les altres.

Es realitzen un nombre fix, n, de proves de Bernouilli

$$X = \text{nombre de vegades que surt E a les n proves} \qquad X \approx \text{Bin}(n,p)$$

¶¶ **La funció densitat de probabilitat** és:

$$f_{\text{Bin}(n,p)}(x) \;:=\; f(k;n,p) = P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k \in \{0,1,2,\ldots,n\} \\ 0 & \text{en cas contrari} \end{cases}$$

on $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$

¶¶ **La funció distribució de probabilitat acumulada** és:

$$F(x;n,p) = P_X(X \le x) = \sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$$

on $[x]$, és *la part entera de* x, és a dir, l'enter més gran que és menor o igual que x.

¶¶ Siga $X \approx \text{Bin}(n,p)$, l'**esperança de X** és: $\mathrm{E}[X] = np$

¶¶ Siga $X \approx \text{Bin}(n,p)$, la **variància de X** és: $\mathrm{Var}[X] = np(1-p)$

Distribució binomial (II)

Exemple 1:
Es tira 2 cops un dau equilibrat. $X =$ nombre de vegades que ix 1. Quina és la probabilitat que no isca cap 1?

$$P_X(0) = \frac{5}{6} \cdot \frac{5}{6}$$

Si fem servir la distribució binomial $X \approx \text{Bin}(2, \frac{1}{6})$, sent: èxit = eixir 1 en un dau equilibrat: $p = 1/6$

$$P_X(0) = \left(\begin{array}{c} 2 \\ 0 \end{array} \right) \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{2-0} = \frac{2!}{0!(2-0)!} 1 \left(\frac{5}{6}\right)^2 = \left(\frac{5}{6}\right)^2$$

Calcular $P_X(1)$, $P_X(2)$.

Exemple 2:
Es llança 5 vegades un dau equilibrat. $X =$ nombre de vegades que ix 1, $X \approx \text{Bin}(5, \frac{1}{6})$

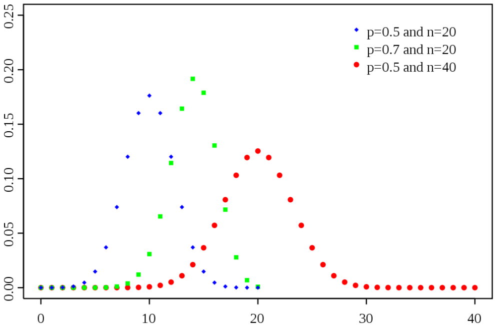
$$P_X(3) = \left(\begin{array}{c} 5 \\ 3 \end{array} \right) \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{5-3}$$

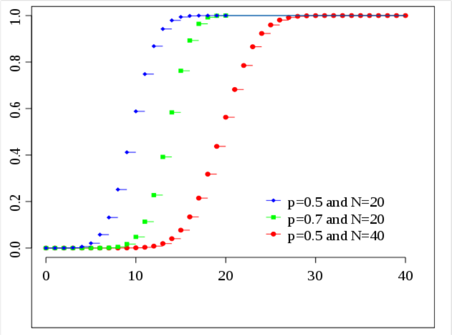
Exemple 3:
Una cadena de muntatge fa 10.000 objectes al dia. El fabricant té constància que un 2 % dels objectes poden ser defectuosos. Un controlador de qualitat avalua 50 objectes al dia. Quina és la probabilitat que trobe 3 objectes defectuosos? En aquest cas tenim una $X \approx \text{Bin}(50, 0.02)$ i la probabilitat és:

$$P_X(3) = \binom{50}{3}(0.02)^3(1-0.02)^{50-3} \simeq 0.061 \Rightarrow 6.1\,\%$$

Distribució binomial (III)

Exemple 4:





Ordres de Matlab:

- `binopdf(X,n,p)`: calcula la *funció de distribució binomial* per al valor `X` i els paràmetres `n` i `p`, amb `n`$\in \mathbb{N}$ i `p` $\in [0,1]$. És a dir,

  $$\texttt{binopdf(X,n,p)} = \binom{n}{X} p^X (1-p)^{n-X}$$

- `binocdf(X,n,p)`: calcula la *funció de distribució de probabilitat acumulada binomial*. El resultat és la probabilitat d'observar fins a `X` successos en `n` proves independents. És a dir,

  $$\texttt{binocdf(X,n,p)} = \sum_{k=0}^{X} \binom{n}{k} p^k (1-p)^{n-k}$$

Distribució geomètrica

$X =$ nombre de vegades que cal repetir una prova de Bernouilli per a obtenir el primer èxit,

$$S_X = \{1, 2, 3, 4, \ldots\ldots\}$$

$$X \quad \approx \quad \mathcal{G}(p), \qquad P_X(k) = p(1-p)^{k-1}$$

Exemple:

1) $X =$ nombre de vegades que cal llançar una moneda perquè isca cara. $p = 0.5$
$P(X = 5) = .5(.5)^4 = 0.0313$, $P(X = 2) = .5(.5) = 0.25$

2) $X =$ nombre de vegades que cal llançar un dau perquè isca 1. $p = 1/6$
$P(X = 5) = \frac{1}{6}(\frac{5}{6})^4 = 0.08$, $P(X = 2) = \frac{1}{6}(\frac{5}{6}) = .1389$

$$P(X \leq 5) = \sum_{i=1}^{5} P(X = i) = \sum_{i=1}^{5} \frac{1}{6}\left(\frac{5}{6}\right)^{i-1} = \frac{1}{6}\sum_{i=1}^{5}\left(\frac{5}{6}\right)^{i-1} = 0.5981$$

¶¶ Siga $X \approx \mathcal{G}(p)$, l'**esperança de X** és: $\mathrm{E}[X] = 1/p$

¶¶ Siga $X \approx \mathcal{G}(p)$, la **variança de X** és: $\mathrm{Var}[X] = \dfrac{1-p}{p^2}$

Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Definicions

Variables aleatòries discretes: Models per a dades que tenen una quantitat numerable de resultats amb probabilitat no nul·la.

¶¶ **Variables aleatòries contínues**: Models per a registres de dades contínues.

Exemples:

- Contingut de glucosa en una solució.
- Mesura del diàmetre interior d'un rodament.
- Mesura de la intensitat de corrent elèctric.

¶¶ Per a una variable aleatòria contínua, $P(X = x) = 0$, però

$$P(X \leq x) = F_X(x)$$

on $F_X(x)$ és la funció de distribució de la variable aleatòria.

Per a una variable aleatòria contínua, $F_X(x)$ és una **funció contínua** $\Longrightarrow$ SENSE salts.

¶¶ Definició: Una variable aleatòria és **(absolutament) contínua** si existeix una funció

$$f : \mathbb{R} \rightarrow \mathbb{R}^+ \qquad \text{tal que:} \qquad F_X(x) = \int_{-\infty}^{x} f(t)dt$$

La funció $f(t)$ és la funció de densitat de probabilitat de la variable aleatòria X.

$$P(X \leq a) = F_X(a) = \int_{-\infty}^{a} f(t)dt$$

$$P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_{a}^{b} f(t)dt, \forall a \leq b$$

Funcions densitat de probabilitat (I)

¶¶ Funció de densitat $f : \mathbb{R} \quad \rightarrow \quad \mathbb{R}^+$ tal que

- $f(x) \geq 0, \quad \forall x \in \mathbb{R}$
- $\displaystyle\int_{-\infty}^{\infty} f(t)dt = 1$

Observeu que $F(x) = \int_{-\infty}^{x} f(t)dt$ és una **primitiva** de $f(t)$, és a dir $F'(x) = f(x)$.

Exemple: $f(x) = \alpha \mathbb{I}_{[0,1]}(x) = \begin{cases} \alpha & x \in [0,1] \\ 0 & \text{en cas contrari} \end{cases}$

- Què ha de valdre α perquè $f(x)$ siga una funció de densitat de probabilitat?
- Calcula $F(x)$
- Què val $P(X > 0.5)$?

¶¶ Distribució uniforme contínua $X \approx U[a,b]$: La seua funció de densitat és

$$f(x) = \frac{1}{b-a}\mathbb{I}_{[ab]}(x), \qquad F(x) = \begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & x \geq b \end{cases}$$

Funcions densitat de probabilitat (II)

¶¶ Distribució exponencial $X \approx \text{Exp}(\lambda)$:

$$f(x) = \lambda e^{-\lambda x} \mathbb{I}_{[0,+\infty]}, \quad F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

¶¶ Distribució NORMAL: Es diu que una variable aleatòria X té una **distribució normal estàndard** (o **distribució gaussiana estàndard**), i s'anota com $X \approx \mathcal{N}(0,1)$, si la seua funció de densitat és (vegeu secció següent):

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad F(x) =!$$

¶¶ Esperança i variància d'una variable aleatòria contínua:
Si X és una variable aleatòria contínua amb funció de densitat $f(x)$, llavors

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad Var(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx$$

Exercici 1:

$X \approx U[a,b]$: $\qquad E(X) = \dfrac{a+b}{2} \qquad \text{Var}(X) = \dfrac{(b-a)^2}{12}$.

$X \approx Exp(\lambda)$: $\qquad E(X) = \dfrac{1}{\lambda} \qquad \text{Var}(X) = \dfrac{1}{\lambda^2}$.

$X \approx \mathcal{N}(0,1)$: $\qquad E(X) = 0, \qquad \text{Var}(X) = 1$

Continguts

1. Introducció
2. Variables aleatòries discretes: Definicions
3. Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat
4. Variables aleatòries discretes: Esperança i variància
5. Variables aleatòries discretes: Funcions densitat de probabilitat
6. Variables aleatòries contínues: Funcions densitat de probabilitat
7. Variables aleatòries contínues: Distribució de Gauss
8. Inferència

Distribució gaussiana (I)

¶¶ Es diu que una variable aleatòria contínua té una distribució **normal estàndard (gaussiana)** si la seua funció de densitat de probabilitat és

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

¶ No existeix una fórmula tancada per a la funció de distribució

$$F(x) = \int_{-\infty}^{x} f(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\left(-\frac{1}{2}z^2\right)} dz$$

¶¶ **L'esperança** d'aquesta variable aleatòria és:

$$E(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{\left(-\frac{1}{2}x^2\right)} dx = 0$$

¶¶ **La variança** d'aquesta variable aleatòria és:

$$Var(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{\left(-\frac{1}{2}x^2\right)} dx = 1$$

¶ Aquesta variable aleatòria s'indica per $\mathbf{X} \approx \mathcal{N}(\mathbf{0}, \mathbf{1})$

¶¶ Es diu que una variable aleatòria contínua té una distribució **gaussiana de mitjana μ i variància σ^2,** $X \approx \mathcal{N}(\mu, \sigma)$ si la seua funció de densitat de probabilitat és

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Distribució gaussiana (II): Propietats

MOLT IMPORTANT: **ESTANDARDITZACIÓ**

1) $E(X) = \mu, \qquad Var(X) = \sigma^2$

2) $X \approx \mathcal{N}(\mu, \sigma) \quad \Longleftrightarrow \quad Z = \frac{X-\mu}{\sigma} \approx \mathcal{N}(0,1)$

¶¶ **Propietats** de la distribució de Gauss:

1. És simètrica respecte de la mitjana.
2. La funció de densitat té punts d'inflexió en $\mu \pm \sigma$.
3. La funció de densitat tendeix asimptòticament de 0 quan $x \to \pm\infty$.
4. $P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$

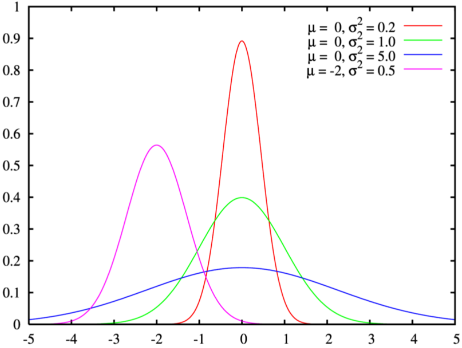
En $[\mu - 2\sigma, \mu + 2\sigma]$ es troba el 95.5 % de l'àrea sota $f(x)$.
En $[\mu - 3\sigma, \mu + 3\sigma]$ es troba el 99.7 %.

¶¶ Moltes variables aleatòries segueixen una **distribució gaussiana**: pesos, alçades, els errors accidentals o aleatoris que es presenten en assaigs de laboratori i observacions experimentals, la descripció del camí aleatori (en física estadística), etc.

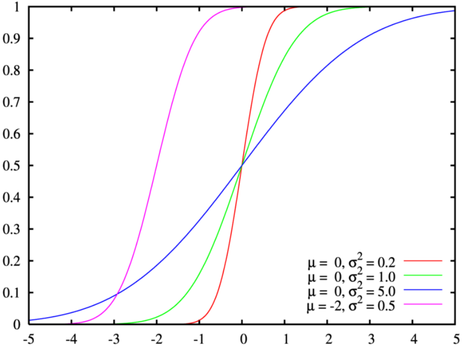
¶¶ Per a valors grans de n, la distribució binomial s'aproxima a la de Gauss.

Distribució gaussiana (III): Exemples

Funció densitat de probabilitat gaussiana



Funció distribució de probabilitat gaussiana integrada



Continguts

1 Introducció

2 Variables aleatòries discretes: Definicions

3 Variables aleatòries discretes: Llei de probabilitat i funció distribució de probabilitat

4 Variables aleatòries discretes: Esperança i variància

5 Variables aleatòries discretes: Funcions densitat de probabilitat

6 Variables aleatòries contínues: Funcions densitat de probabilitat

7 Variables aleatòries contínues: Distribució de Gauss

8 Inferència

Inferència estadística

La inferència és la part de l'estadística que es dedica a fer prediccions sobre valors de paràmetres poblacionals (desconeguts) a partir de les dades d'una mostra. Inclou aquells procediments que permeten prendre decisions en problemes de naturalesa aleatòria.

Objectiu: Obtenir informació d'una població a la qual només es pot accedir parcialment, és a dir, una població de la qual podem (o volem) analitzar només alguns individus. Es pretén estimar paràmetres a partir de quantitats que es dedueixen de valors mostrals i que es denominen **estadístics**.

Com: Un cert model probabilístic es considera apropiat per a descriure el fenomen aleatori objecte d'estudi. Es vol esbrinar, algun o alguns dels paràmetres que defineixen la funció de distribució del model, a partir de dades (experimentals) basades en mostres de la població estudiada.

Exemple: Determineu l'alçada mitjana dels xiquets de 10 anys residents a la ciutat de València, a partir de les alçades d'una mostra formada per 400 xiquets.

L'alçada mitjana obtinguda a partir de les dades de la mostra és un estimador de la mitjana de la població. Necessitem instruments per a mesurar la seua fiabilitat.

Inferència estadística

Els mètodes per a realitzar inferència es poden classificar en dos tipus:

1. es pot predir o estimar el valor del paràmetre amb una probabilitat determinada d'encertar, o bé
2. poden prendre decisions sobre el valor del paràmetre i posteriorment comprovar la seua validesa.

És a dir, els mètodes estadístics proporcionen, a més del valor estimat del paràmetre, *una mesura de la seua fiabilitat*.

Situació habitual: Es coneix quin és el model probabilístic adequat, però falta informació sobre algun o alguns dels paràmetres que caracteritzen la seua funció de distribució. L'estimació del valor d'un paràmetre poblacional desconegut pot ser:

- Estimació puntual: determinació de valors que representen estimacions raonables del paràmetre en qüestió, a partir d'una mostra de la població.
- Estimació per intervals de confiança: determinació d'intervals en què s'espera trobar el paràmetre amb un cert nivell de confiança.

Inferència estadística

Es desitja obtenir estimacions raonables d'un (o diversos) dels paràmetres que defineixen el model probabilístic, a partir de les dades d'una mostra de la població.

Exemple:
X $=$ alçada d'un xiquet de 10 anys, triat a l'atzar $\approx \mathcal{N}(\mu, \sigma)$. Es vol estimar μ i/o σ a partir d'una mostra de la població.

Una vegada realitzada l'anterior estimació s'ha d'estudiar si és o no acceptable a partir de les dades mostrals de què es dispose. Per a això s'utilitza el **contrast d'hipòtesis**, que ens permet acceptar o rebutjar una certa hipòtesi sobre el valor d'un paràmetre.

Alguns exemples

- Estimar la mitjana de les estatures dels xiquets de 10 anys a la ciutat de València

  $X =$ alçada d'un xiquet de 10 anys que viu a València.
  Determineu $E(X)$, el valor esperat per a l'alçada o mitjana poblacional.

Com podem fer això?

Alguns exemples

- Estimar la mitjana de les estatures dels xiquets de 10 anys a la ciutat de València

  $X =$ alçada d'un xiquet de 10 anys que viu a València.
  Determineu $E(X)$, el valor esperat per a l'alçada o mitjana poblacional.

Com podem fer això?

- Prenem una *mostra* de n xiquets i en mesurem les alçades.
- Esperem que la mitjana de les alçades de la mostra represente la mitjana de la població, $E(X)$?

$\Longrightarrow$ Necessitem un model probabilístic per a X.

$\Longrightarrow$ Necessitem entendre els rudiments de la teoria del mostreig aleatori.

Teoria elemental del mostreig

Un procediment de mostreig no és més que un procés que permet seleccionar elements (mostres) en una població.

Mostreig aleatori simple: Procés de selecció d'elements en una població, de manera que cada element té idèntica probabilitat de ser elegit en cada extracció.

En poblacions finites, un mostreig aleatori simple és equivalent a un procés d'extraccions amb reemplaçament.

Exemple: Se sap que l'alçada dels xiquets de 10 anys de València segueix una distribució normal de mitjana μ i variància σ^2.
$X =$ alçada d'un xiquet de 10 anys triat a l'atzar $\approx \mathcal{N}(\mu, \sigma)$.

Escollim a l'atzar 400 xiquets triats de la població {xiquets de 10 anys de València}.

$X_i =$ alçada del xiquet i a la mostra. $X_i \approx \mathcal{N}(\mu, \sigma)$, $i = 1, \ldots, n$
$(X_1, X_2, \ldots. X_{400})$ és una mostra aleatòria simple de X de mida 400.

Teoria elemental del mostreig: Estimació de l'esperança

Definició: Un mostreig aleatori simple (*m.a.s.*) de mida n per a una variable aleatòria X és un vector aleatori $(X_1, X_2, \ldots, X_n)$ on X_i, $i = 1, \ldots, n$ són variables aleatòries independents, idènticament distribuïdes (i.i.d.) que segueixen el mateix model probabilístic que X.

Cada component d'un *m.a.s.* representa els possibles valors de la component i-èsima d'una mostra aleatòria de grandària n de població.

$X=$ alçada d'1 xiquet de 10 anys triat a l'atzar $\approx \mathcal{N}(\mu, \sigma)$.
μ és la mitjana poblacional i σ^2 la variància poblacional

$X_i =$ alçada del xiquet i en la mostra. $X_i \approx \mathcal{N}(\mu, \sigma)$, $i = 1, \ldots, n$
$(X_1, X_2, \ldots. X_{400})$ és un m.a.s de mida 400 de X.

IMPORTANT: Estimació de l'esperança

La mitjana mostral $\quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad$ és una estimació per μ

Estimació per intervals de confiança

$X \approx \mathcal{N}(\mu, \sigma)$. $(X_1, X_2, \dots. X_n)$ m.a.s de mida n de X. $X_i \approx \mathcal{N}(\mu, \sigma)$

$$\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

és una v. a., que és un estimador per a μ.

- El valor d'un estimador per a una mostra concreta és una estimació.
- Una estimació depèn de la mostra poblacional escollida, de manera que si s'extrauen dues mostres diferents, les estimacions seran diferents.
- És desitjable proporcionar una mesura de la precisió de l'estimació del paràmetre a partir de l'estadístic (estimador) considerat.

Els intervals de confiança proporcionen aquesta mesura.

Definició:

Siga X una v. a. que depèn d'un paràmetre desconegut θ. Siga $(X_1, X_2, \dots, X_n)$ un m.a.s. de X. Anomenarem interval de confiança de nivell $1-\alpha$, $\alpha \in (0,1)$, a un interval $[L_1, L_2]$, els extrems del qual depenen del m.a.s., i tal que

$$P(L_1 \leq \theta \leq L_2) = 1 - \alpha$$

Estimació per intervals de confiança

Siga $(X_1, X_2, ..., X_n)$ un m.a.s. de mida n d'una variable aleatòria X, i θ un parametre poblacional de X. El problema de l'estimació per interval de confiança per a un paràmetre θ consisteix a determinar un interval $[L_1, L_2]$ tal que

$$P(L_1 \leq \theta \leq L_2) = 1 - \alpha$$

per a una certa funció de densitat de probabilitat P.

De tots els possibles intervals que compleixen la condició anterior, es tria el de menor amplada que coincidisca amb el de major densitat de probabilitat, i que estiga centrat en la mitjana mostral.

El càlcul pràctic d'intervals de confiança per a un paràmetre poblacional desconegut es basa en:

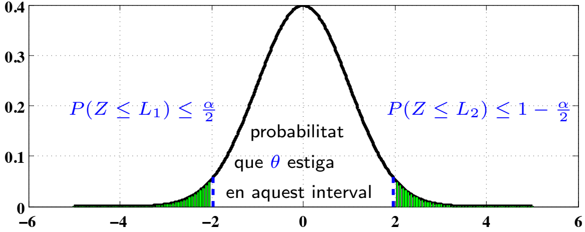
1. Un estimador no esbiaixat del paràmetre a estimar.
2. Una v. a. de distribució coneguda, que anomenarem Z, construïda a partir de l'estimador obtingut del m.a.s.
3. La funció de distribució de probabilitat per a Z és la que s'utilitzarà per a la determinació de l'interval de confiança.

Estimació per intervals de confiança

Per a determinar l'interval de confiança més xicotet possible que centrat en la mitjana complisca els requisits exigits cal avaluar L_1 i L_2 per als quals

$$P(Z \leq L_1) \leq \frac{\alpha}{2}, \qquad P(Z \leq L_2) \leq 1 - \frac{\alpha}{2}.$$

Exemple si la v. a. es distribueix segons $\mathcal{N}(0,1)$.



Nota: si la funció de densitat de probabilitat de Z és simètrica respecte a 0 (p. ex., la normal tipificada, o t-Student), es compleix que

$$-L_1 = L_2 \equiv -z^*$$

Habitualment, el que més ens interessarà és estimar els intervals de confiança per a la mitjana d'una població, distribuïda d'acord a una funció normal, amb un nivell de confiança $1-\alpha$, distingint dues possibles situacions:

1. La variància poblacional σ^2 és coneguda.
2. La variància poblacional σ^2 és desconeguda.

IC per μ, coneixent σ

Siga $X \approx \mathcal{N}(\mu, \sigma)$ amb μ desconeguda i σ coneguda.

Volem estimar μ amb un nivell de confiança $1 - \alpha$ a partir d'un m.a.s.:

$(X_1, \ldots, X_n)$, X_i i.i.d, $\approx \mathcal{N}(\mu, \sigma)$.
(El nivell de confiança més habitual és el 95 %)

Considerem la mitjana mostral $$\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

Com $X_i \approx \mathcal{N}(\mu, \sigma) \to \boxed{\bar{X} \approx \mathcal{N}(\mu, \sigma/\sqrt{n})}$.

Estandarditzem $$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx \mathcal{N}(0, 1)$$

Z és una v. a. amb distribució coneguda.

IC per μ, coneixent σ

Calculem un interval centrat $[-z^*, z^*]$ tal que

$$P(-z^* \leq Z \leq z^*) = 1 - \alpha$$

El valor de z^* es calcula utilitzant Matlab/taules (de la normal estàndard) de la següent manera:

$$-z^* \leq Z \leq z^* \quad \equiv \quad -z^* \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z^* \quad \equiv \quad \bar{X} - z^* \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z^* \frac{\sigma}{\sqrt{n}}$$

$$1 - \alpha = P(-z^* \leq Z \leq z^*) = P(\bar{X} - z^* \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} - z^* \frac{\sigma}{\sqrt{n}})$$

Per tant, l'interval

$$\left[\bar{X} - z^* \frac{\sigma}{\sqrt{n}}, \bar{X} + z^* \frac{\sigma}{\sqrt{n}}\right]$$

on z^* es calcula de manera que $P(Z \leq z^*) = 1 - \alpha/2$, és un IC per μ amb nivell de confiança $1 - \alpha$.

α es diu nivell de significació. Si es canvia α, es canvia z^*.

IC per μ, coneixent σ

Resumint:
$X \approx \mathcal{N}(\mu, \sigma)$, $(X_1, \ldots, X_n)$ m.a.s. de X.

Per a calcular el IC per μ amb nivell de confiança $1-\alpha$, determinem un interval centrat que continga el $(1-\alpha)100\,\%$ de la probabilitat de la v. a. Z, és a dir, determinem z^* tal que

$$P(-z^* \leq Z \leq z^*) = 1-\alpha$$

El nivell de confiança està directament relacionat amb el valor z^*.
per a $1-\alpha = 0.95$ obtenim $z^* = 1.96$
per a $1-\alpha = 0.99$, s'obté $z^* = 2.58$
per a qualsevol altre nivell de significació, podem usar: `norminv(1-`α`/2,0,1)`

L'interval

$$\left[\bar{X} - z^* \frac{\sigma}{\sqrt{n}}, \bar{X} + z^* \frac{\sigma}{\sqrt{n}}\right]$$

és un interval de confiança amb nivell de confiança $1-\alpha$ per μ.

Exemple

S'ha registrat el valor (en kg) de la reducció del pes de cadascun dels pacients d'una mostra escollida a l'atzar en una determinada població, després d'una setmana de tractament. La mitjana dels 16 valors obtinguts és de 3.42 kg. Suposant que la pèrdua de pes és una v. a. amb una distribució $\mathcal{N}(\mu, \sigma)$, construïu un interval de confiança del 95 % per al valor mitjà poblacional de la reducció de pes després d'una setmana de tractament

1. en el cas en què $\sigma = 0.68$ kg
   `norminv(.975,0,1)=1.96` `Sol: IC=` $[3.0868, 3.7532]$

2. en el cas en què σ siga desconeguda, però s'haja mesurat la variància poblacional i s'haguera obtingut que $S = 0.68$ kg
   La solució NO és $[3.0868, 3.7532]$

Per què no es pot calcular l'IC quan no es coneix σ?

IC per μ, si no es coneix σ

Si σ és desconegut, $\bar{X} \pm z^* \frac{\sigma}{\sqrt{n}}$ no es pot calcular.

Si no es coneix σ, cal utilitzar estimadors per a σ.

- El millor estimador per a la variància σ^2 NO és la variància mostral

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- El millor estimador per a la variància σ^2 és la quasi-variància mostral S_c^2

$$\boxed{S_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

Noteu que $S_c^2 = \frac{n}{n-1}S^2$

- Si n és gran ($n > 30$) i σ^2 és desconeguda:

$$Z = \frac{\bar{X} - \mu}{S_c/\sqrt{n}} \approx \mathcal{N}(0,1)$$

IC per μ amb σ desconegut

Per a construir un IC per μ en poblacions que segueixen el model normal, a partir d'una mostra amb n gran $(n > 30)$ es considera

$$Z = \frac{\bar{X} - \mu}{S_c/\sqrt{n}} \approx \mathcal{N}(0,1)$$

i es construeix l'IC a partir d'aquesta variable. L'IC amb nivell de confiança $(1-\alpha)$ és

$$[\bar{X} - z^* \frac{S_c}{\sqrt{n}}, \bar{X} + z^* \frac{S_c}{\sqrt{n}}]$$

on z^* satisfà $P(Z \leq z^*) = 1 - \alpha/2$.

Exercici 2: Un fabricant de fils desitja estimar la tensió mitjana de ruptura de la seua producció. Per a això mesura la tensió fins a la ruptura d'una mostra de 100 fils. Per a aquesta mostra s'obté $\bar{X} = 110.8$, $S_c = 10.2$. Calculeu intervals de confiança al 95 % i al 99 % per a la tensió mitjana, suposant normalitat.

IC per μ amb σ desconegut

Exemple:

La següent sèrie de valors representa el temps de funcionament (en hores) sense recàrrega d'un dispositiu que estem dissenyant. Determineu l'interval de confiança (al 99 %) que podeu donar en el manual de l'esmentat dispositiu quan el tragueu al mercat.

72.419	74.604	74.223	73.085	73.696	74.240
73.055	76.124	74.508	73.956	74.500	73.857
74.986	74.992	74.470	73.155	74.502	75.141
74.509	73.788	74.206	73.449	74.622	73.197
71.939	75.007	74.228	73.472	74.959	72.802

Amb aquests valors obtenim:

$$\bar{X} = 74.056 \simeq \mu$$
$$S_c^2 = 0.8212$$
$$z^* = 2.58$$
$$1 - \alpha = 0.99$$

Per tant, amb una probabilitat del 99 %, el valor de la mitjana d'hores de funcionament del dispositu es trobarà en l'interval:

$$\left[74.056 - 2.58 \cdot \frac{\sqrt{0.8212}}{\sqrt{30}}, 74.056 - 2.58 \cdot \frac{\sqrt{0.8212}}{\sqrt{30}}\right] = [73.6003, 74.5124]$$

IC per μ amb σ desconegut

Si la mostra no és gran $n \leq 30$, llavors

$$Z = \frac{\bar{X} - \mu}{S_c/\sqrt{n}}$$

no es comporta com una v. a. $\mathcal{N}(0,1)$
sinó com una t de Student, amb $n-1$ graus de llibertat. t-Student vs. Normal

Característiques del model t de Student

- Una v. a. X que segueix el model t de Student es caracteritza per un paràmetre: $m =$ els graus de llibertat $X \approx t_m$
- La funció de densitat de probabilitat $f(x)$ d'una $X \approx t_m$ té forma de campana invertida, però és més ampla que la normal estàndard.
- $X \approx t_m \rightarrow E(X) = 0$ i $f(x)$ és simètrica respecte de 0, el seu valor mitjà.
- Quan m és gran, la $f(x)$ d'una v. a. t_m i la d'una $\mathcal{N}(0,1)$ són pràcticament iguals.

  $m > 30$ **se sol considerar 'gran'**.

IC per μ amb σ desconegut

Per a construir IC per μ en poblacions que segueixen el model normal, a partir d'una mostra amb $n \leq 30$ es considera

$$Z = \frac{\bar{X} - \mu}{S_c/\sqrt{n}} \approx t_{n-1}$$

i es construeix l'IC a partir d'aquesta variable.

L'IC amb nivell de confiança $(1-\alpha)$ és

$$[\bar{X} - t^* \frac{S_c}{\sqrt{n}}, \bar{X} + t^* \frac{S_c}{\sqrt{n}}]$$

t^* satisfà $P(Z \leq t^*) = 1 - \alpha/2$.

El valor de t^* es busca en les taules de la t -student (o usant Matlab), segons la mida de la mostra.

IC per μ amb σ desconegut

Exemple:
S'ha registrat el valor (en kg) de la reducció del pes de cada un dels pacients d'una mostra escollida a l'atzar en una determinada població, després d'una setmana de tractament. La mitjana dels 16 valors obtinguts és de 3.42 kg, i s'ha mesurat la variància poblacional obtenint $S_c = 0.68$ kg. Suposant que la pèrdua de pes és una v. a. amb una distribució $\mathcal{N}(\mu, \sigma)$, construïu un interval de confiança del 95 % per al valor mitjà poblacional de la reducció de pes després d'una setmana de tractament.

Solució:
En Matlab farem servir la instrucció `tinv(1-`α`/2,m)` per a calcular el valor de t^*, de manera que, com que en el nostre cas

$$1 - \alpha = 0.95 \Longrightarrow \alpha = 0.05 \Longrightarrow 1 - \alpha/2 = 0.975,$$

`tinv(0.975,16)=2.1199` $\equiv t^*$,

per tant $\mathtt{IC}_{95\,\%} = [3.42 - 2.1199\frac{0.68}{\sqrt{16}},\ 3.42 + 2.1199\frac{0.68}{\sqrt{16}}] = [3.0596, 3.7804]$

Matemàtiques III:
7. Optimització convexa bàsica

Miguel Ángel Aloy Torás

Departament d'Astronomia i Astrofísica
Tutories: Dilluns de 10 h a 13 h
Despatx 1.54. Edifici Jeroni Muñoz.

miguel.a.aloy@uv.es

Curs 2015-2016

Continguts

1 Introducció

2 Resolució de problemes d'optimització

3 Programació lineal

Bibliografia:

Boyd, S. & Vandernberghe, L., Convex Optimization, http://stanford.edu/∼boyd/cvxbook

Continguts

1 Introducció

2 Resolució de problemes d'optimització

3 Programació lineal

Optimització matemàtica

El problema de l'optimització matemàtica

$$\begin{array}{ll} \text{minitmizar} & f_0(x) \\ \text{subjecte a} & f_i(x) \leq b_i, \quad i = 1, \ldots, m \end{array}$$

- $x = (x_1, \ldots, x_n)$: variables d'optimització.
- $f_0 : \mathbb{R}^n \longrightarrow \mathbb{R}$: funció objectiu, el domini de la qual és $\chi \subset \mathbb{R}^n$
- $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}, \ \ i = 1, \ldots, m$: restriccions.

La **solució òptima** x^* es correspon amb la que proporciona el menor valor de f_0 d'entre tots els vectors que satisfan les restriccions, és a dir,

$$f_0(x^*) \leq f_0(x), \forall x \in \chi.$$

Exemples

Optitmizació de carteres

- Variables: quantitats invertides en diferents accions.
- Restriccions: pressupost, max./min. inversió per acció, mínim retorn.
- Objectiu: risc conjunt o variància del retorn.

Dimensionament de dispositius en circuits electrònics

- Variables: amplària i longitud dels dispositius.
- Restriccions: límits de producció, requeriments temporals, àrea màxima.
- Objectiu: consum elèctric.

Ajust de dades

- Variables: paràmetres del model.
- Restriccions: informació prèvia, límits sobre els paràmetres.
- Objectiu: mesura del desajust o predicció de l'error comès en comparar dades observades i valors predits pel model.

Continguts

1 Introducció

2 Resolució de problemes d'optimització

3 Programació lineal

Resolució de problemes d'optimització

Solució general del problema d'optimització

- Molt difícil de resoldre.
- Els mètodes inclouen algun tipus de compromís, p. ex. temps de càlcul molt llarg o la possibilitat que no sempre es puga trobar la solució.

Excepcions: certes classes de problemes poden ser resolts de forma eficient i fiable (fins i tot en casos grans amb centenars o milers de variables i restriccions).

- Problemes de mínims quadrats.
- Problemes de programació lineal.
- Problemes d'optimització convexa.

Mínims quadrats (MQ)

minimitzar $\quad f_0(x) = \|Ax - b\|_2^2 = \sum_{i=1}^k (a_i^T x - b_i)^2$
sense restriccions

$A \in \mathbb{R}^{k\times n}$ (amb $k \geq n$), a_i^T són les files de A, i el vector $x \in \mathbb{R}^n$ és la variable d'optimització.

Resolució de problemes de mínims quadrats

- Solució analítica: $x^* = (A^T A)^{-1} A^T b$.
- Hi ha algoritmes i programari molt eficient per a resoldre'ls.
- El temps de càlcul és proporcional a $n^2 k$ ($A \in \mathbb{R}^{k\times n}$); excepte si A té certes estructures.
- Es tracta d'una tecnologia madura.

Ús de mínims quadrats

- Els problemes de MQ són fàcils de reconèixer (se n'han vist alguns en el tema 1).
- Hi ha unes poques tècniques estàndard que incrementen la flexibilitat (p. ex., incloent-hi pesos, afegint-hi termes de regularització, etc.)

¶ Problemes de MQ: cas particular de problemes d'optimització convexa.

Programació lineal (PL)

$$\begin{array}{ll} \text{minimitzar} & f_0(x) = c^T x \\ \text{subjecte a} & f_i(x) = a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

$c, a_1, \ldots, a_m \in \mathbb{R}^n$ (vectors) i $b_1, \ldots, b_m \in \mathbb{R}$ (paràmetres escalars).

Resolució de problemes de programació lineal

- No hi ha (en general) solució analítica (diferència dels mètodes MQ).
- Però existeixen algoritmes i programari molt eficient per a resoldre'ls (p. ex., mètode Simplex, mètode del punt interior, etc.).
- $t_{\text{càlcul}} \propto n^2 m$ (suposant $m > n$ i calculant el temps usant el mètode del punt interior); excepte si A té certes estructures.
- Es tracta d'una tecnologia madura (tot i que, potser, no al nivell dels MQ).

Ús de la programació lineal

- Els problemes de PL no són tan fàcils de reconèixer com els de MQ. Vegeu el *poblema d'aproximació de Txebixev* (Boyd & Vanderberghe, Eq. 1.6)
- Hi ha unes poques tècniques estàndard que converteixen problemes d'optimització en problemes de PL (p. ex., problemes que depenguen de les normes l_1 (p. ex., Txebixev) o l_∞, funcions lineals a trossos, etc.).

¶ Problemes de PL: cas particular de problemes d'optimització convexa.

Problemes d'optimització convexa (OC) (I)

$$\begin{array}{ll} \text{minimitzar} & f_0(x) \\ \text{sujecte a} & f_i(x) \leq b_i, \quad i = 1, \ldots, m \end{array}$$

les funcions $f_0, \ldots, f_m : \mathbb{R}^n \Longrightarrow \mathbb{R}$ són convexes.

- **Les funcions objectiu i restricció són totes convexes:**

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y)$$

$\forall x, y \in \mathbb{R}^n$ y $\forall \alpha, \beta \in \mathbb{R},\ \alpha + \beta = 1,\ \alpha \geq 0,\ \beta \geq 0$

IMPORTANT

Suposem que $\exists f''(x),\ \forall x \in [a, b]$:

$$f''(x) \geq 0 \Longleftrightarrow f(x) \text{ és convexa } \forall x \in [a, b]$$

- Inclou MQ i PL com a casos especials. En el cas de PL, les funcions són lineals:

$$f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y).$$

- La convexitat és un concepte més general que el de la linealitat: una desigualtat reemplaça el cas més restrictiu de la igualtat de la definició de funció lineal, i a més, no qualsevol valor de α i β és vàlid (només certes combinacions).

Problemes d'optimització convexa (OC) (II)

Resolució de problemes d'optimització convexa

- No hi ha (en general) solució analítica.
- Però hi ha algoritmes i programari molt eficient per a resoldre'ls.
- $t_{\text{càlcul}} \propto \max\{n^3, n^2 m, F\}$ (suposant $m > n$ i calculant el temps usant el mètode del punt interior); excepte si A té certes estructures. F és el cost d'avaluar les f_i 's i les seues primeres i segones derivades.
- Es tracta gairebé d'una tecnologia (és més difícil plantejar el problema en la forma adequada que resoldre'l).

Ús de l'optimització convexa

- Els problemes d'OC són difícils de reconèixer (sovint).
- Hi ha multitud de tècniques que converteixen problemes d'optimització en problemes d'OC (però no hi ha procediments *estàndard*),
- Sorprenentment, molts problemes es poden resoldre via OC.

Continguts

1 Introducció

2 Resolució de problemes d'optimització

3 Programació lineal

Programació lineal

$$\begin{array}{ll} \text{minimitzar} & f_0(x) = c^T x \\ \text{sujecte a} & f_i(x) = a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

$c, a_1, \ldots, a_m \in \mathbb{R}^n$ (vectors) i $b_1, \ldots, b_m \in \mathbb{R}$ (paràmetres escalars).

És un cas particular d'OC en què les f_i són totes lineals.

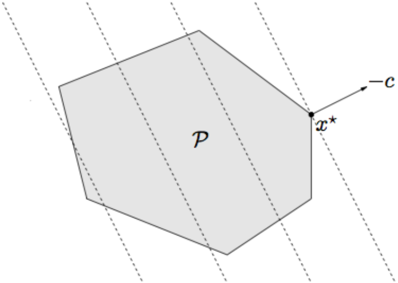


Figure 4.4 Geometric interpretation of an LP. The feasible set $\mathcal{P}$, which is a polyhedron, is shaded. The objective $c^T x$ is linear, so its level curves are hyperplanes orthogonal to c (shown as dashed lines). The point $x^\star$ is optimal; it is the point in $\mathcal{P}$ as far as possible in the direction $-c$.

Programació lineal

$$\begin{array}{ll} \text{minimitzar} & f_0(x) = c^T x \\ \text{sujecte a} & f_i(x) = a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

$c, a_1, \ldots, a_m \in \mathbb{R}^n$ (vectors) i $b_1, \ldots, b_m \in \mathbb{R}$ (paràmetres escalars).

És un cas particular d'OC en què les f_i són totes lineals.

¶ **Exemple:**

Una companyia fabrica i ven dos models de làmpada L_1 i L_2. Per a la seua fabricació es necessita un treball manual de 20 hores per al model L_1 i de 30 hores per al L_2; i un treball de màquina de 25 hores per L_1 i de 10 hores per al L_2. Es disposa per al treball manual de 100 hores al mes i per a la màquina de 80 hores al mes. Sabent que el benefici per unitat és de 15 i 10 euros per L_1 i L_2, respectivament, planifiqueu la producció per a obtenir el màxim benefici.

Programació lineal

¶ **Exemple:**

Plantejament matemàtic del problema:

maximitzar $f_0(L_1, L_2) = 15L_1 + 10L_2$ (beneficis = funció objectiu)

subjecte a
$$
\begin{aligned}
f_1(L_1, L_2) &= 20L_1 + 30L_2 \leq 100 \\
f_2(L_1, L_2) &= 25L_1 + 10L_2 \leq 80 \\
& L_1 \geq 0, \quad L_2 \geq 0
\end{aligned}
$$

Programació lineal

¶ **Exemple:**

Plantejament matemàtic del problema:

maximitzar $f_0(L_1, L_2) = 15L_1 + 10L_2$ (beneficis = funció objectiu)

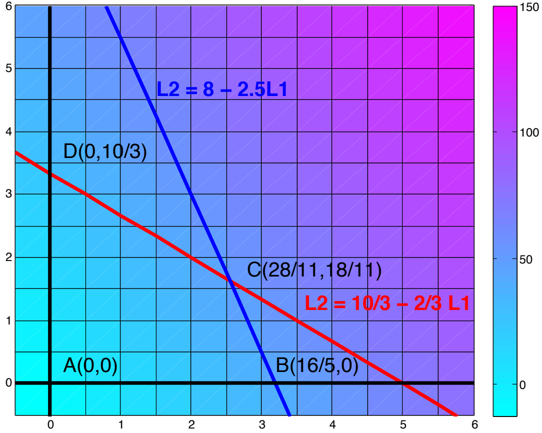
subjecte a
$$\begin{aligned} f_1(L_1, L_2) &= 20L_1 + 30L_2 \leq 100 \\ f_2(L_1, L_2) &= 25L_1 + 10L_2 \leq 80 \\ L_1 \geq 0,& \quad L_2 \geq 0 \end{aligned}$$

¶¶ Veurem dos mètodes diferents per a resoldre aquest tipus de problemes:

1. **Mètode gràfic** (vàlid només per a dues variables)
2. **Mètode Matlab**

Programació lineal: Mètode gràfic

1. Representem gràficament la regió delimitada per les restriccions:



2. Tant la funció objectiu com les restriccions són funcions lineals en $x = (L_1, L_2)$. Per tant, el màxim/mínim de la funció objectiu s'aconsegueix en el contorn del polígon (domini) que determinen les restriccions.

Programació lineal: Mètode gràfic

1. De fet, el màxim de la funció objectiu s'aconsegueix en algun dels vèrtexs del domini A, B, C o D, els quals són:

$$A = (0,0);\ B = (16/5, 0);\ C = (28/11, 18/11);\ D = (0, 10/3)$$

2. La funció objectiu pren en aquests punts els següents valors:

$$f(A) = 0;\ f(B) = 48;\ f(C) = 600/11 \simeq 54.55;\ f(D) = 100/3 \simeq 33.33$$

Per tant, la solució al problema és el punt C:

$$L_1 = 28/11, \quad L_2 = 18/11, \quad f_0(L_1, L_2) = 600/11$$

¶ **Exercici 1:**

$$\begin{array}{ll} \text{maximitzar} & f_0(x_1,x_2) = 2x_1 + 3x_2 \\ \text{sujecte a} & f_1(x_1,x_2) = 5x_1 + 2x_2 \leq 10 \\ & f_2(x_1,x_2) = x_1 + 2x_2 \leq 8 \\ & \qquad x_1 \geq 0, \quad x_2 \geq 0 \end{array}$$

Programació lineal: Mètode gràfic

¶ Exemple:

Tenim una empresa que fabrica dos tipus de dispositius electrònics. Seguim un procediment manual per a soldar els dispositius a la placa i som capaços de muntar, per dia, 2 del primer tipus i 3 del segon. A més la màquina que encapsula les plaques només pot processar (com a molt) 4 dispositius per dia. El benefici que traiem és de 20 EUR per les màquines de primer tipus i 10 EUR per les del segon. Quina producció he de tenir al dia per a maximitzar el benefici?

1. Identifiquem les variables del problema (variables de decisió):
   - Nombre de productes del primer tipus: P_1
   - Nombre de productes del segon tipus: P_2
2. Identifiquem les restriccions del problema:
   - Ací cal tenir en compte que la producció ha de ser positiva, aquest tipus de restricció no s'indica però es dóna per descomptat: $P_1 \geq 0$, $P_2 \geq 0$.
   - Ens diuen que: $P_1 \leq 2$, $P_2 \leq 3$.
     $$\Longrightarrow 0 \leq P_1 \leq 2,\ 0 \leq P_2 \leq 3$$
   - També ens diuen que: $P_1 + P_2 \leq 4$
3. Plantegem la funció objectiu (beneficis): $f_0(P_1, P_2) = 20P_1 + 10P_2$

Programació lineal: Mètode gràfic

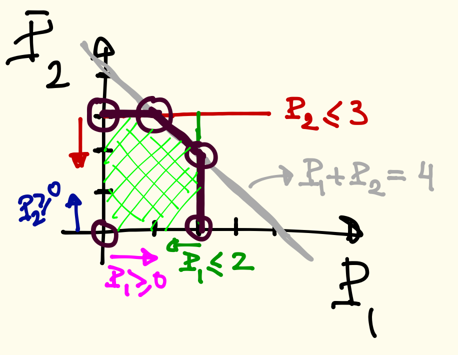
Plantejament matemàtic del problema:

maximitzar $f_0(P_1, P_2) = 20P_1 + 10P_2$ (beneficis = funció objectiu)

subjecte a
$$
\begin{aligned}
f_1(P_1, P_2) &= P_1 + P_2 \leq 4 \\
f_2(P_1, P_2) &= P_1 \leq 2 \\
f_3(P_1, P_2) &= P_2 \leq 3 \\
&P_1 \geq 0, \quad P_2 \geq 0
\end{aligned}
$$

Programació lineal: Mètode gràfic

1. Representem gràficament la regió delimitada per les restriccions:



2. Tant la funció objectiu com les restriccions són funcions lineals en $x = (P_1, P_2)$. Per tant, el màxim/mínim de la funció objectiu s'aconsegueix en el contorn del polígon (domini) que determinen les restriccions.

Programació lineal: Mètode gràfic

1. Els vèrtexs del domini on totes les restriccions es compleixen són:

$$P(0,0); P(2,0); P(0,3), P(1,3) \text{ i } P(2,2).$$

2. Tant la funció objectiu com les restriccions són funcions lineals en $x = (P_1, P_2)$. Per tant, el màxim/mínim de la funció objectiu s'aconsegueix en el contorn del polígon (domini) que determinen les restriccions.
3. El valor de la funció objectiu en els vèrtexs del domini és:

$$f_0(0,0) = 0, f_0(2,0) = 40, f_0(0,3) = 30, f_0(1,3) = 50, f_0(2,2) = 40$$

Solució: He de produir 2 unitats de cada producte per a maximitzar el benefici.

¶ **Exercici 2:** Representeu la funció $15P_1 + 10P_2 = K$, donant-li valors a K de 15, 30, 45. Com són aquestes rectes?

Programació lineal: Resolució amb Matlab

Un problema general de programació lineal pot definir-se com segueix:

$$\begin{array}{ll} \text{minimitzar} & f_0(x) = f^T \cdot x \\ \text{sujecte a} & A \cdot x \leq b \\ & A_{eq} \cdot x = b_{eq} \\ & lb_i \leq x_i \leq ub_i, \quad i = 1, \ldots, n \end{array}$$

on f_0 és la funció objectiu (a minimitzar); $A \in \mathbb{R}^{m \times n}$ i $A_{eq} \in \mathbb{R}^{l \times n}$ són sengles matrius; $x, f, lb, ub \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, i $b_{eq} \in \mathbb{R}^l$ són vectors.

Aquest problema el resol Matlab emprant la funció `linprog`, la qual disposa de multitud d'opcions, la major part de les quals es deixen a la curiositat de l'alumne (fent servir el comandament `help linprog`).

Programació lineal: Resolució amb Matlab

Un problema general de programació lineal pot definir-se com segueix:

$$
\begin{array}{ll}
\text{minimitzar} & f_0(x) = f^T \cdot x \\
\text{sujecte a} & A \cdot x \leq b \\
 & A_{eq} \cdot x = b_{eq} \\
 & lb_i \leq x_i \leq ub_i, \quad i = 1, \ldots, n
\end{array}
$$

on f_0 és la funció objectiu (a minimitzar); $A \in \mathbb{R}^{m \times n}$ i $A_{eq} \in \mathbb{R}^{l \times n}$ són sengles matrius; $x, f, lb, ub \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, i $b_{eq} \in \mathbb{R}^l$ són vectors.

Nosaltres utilitzarem la sintaxi:

```
[x, fval] = linprog(f, A, b, Aeq, beq, lb, ub, x0)
```

Sent:

- `x`: valor del vector solució al problema
- `f`: vector que caracteritza la funció a minimitzar
- `A, b, Aeq, beq, lb, ub`: restriccions del problema
- `x0`: valor inicial de `x`, si fóra necessari
- `fval`: valor de la funció objectiu per a la `x` òptima.

Programació lineal: Resolució amb Matlab

¶ **Exemple:**

maximitzar $f_0(L_1, L_2) = 15L_1 + 10L_2$ (beneficis = funció objectiu)

sujecte a

$$
\begin{aligned}
f_1(L_1, L_2) &= 20L_1 + 30L_2 \leq 100 \\
f_2(L_1, L_2) &= 25L_1 + 10L_2 \leq 80 \\
L_1 \geq 0&, \quad L_2 \geq 0
\end{aligned}
$$

Com que el problema proposat ens demana maximitzar la funció objectiu en lloc de minimitzar (el que fa `linprog`), hem d'agafar

$$-f_0(L_1, L_1) = -15L_1 - 10L_2$$

i procedir amb la minimització d'aquesta funció objectiu.

Amb Matlab farem servir la següent sintaxi:

```
>> f = [-15; -10];
>> A = [20 30; 25 10];
>> b = [100; 80];
>> lb = zeros(2,1);
>> [x, fval] = linprog(f, A, b, [], [], lb, [], [])
```

Programació lineal: Resolució amb Matlab

¶ **Exemple:**

maximitzar $f_0(L_1, L_2) = 15L_1 + 10L_2$ (beneficis = funció objectiu)

sujecte a

$$f_1(L_1, L_2) = 20L_1 + 30L_2 \leq 100$$
$$f_2(L_1, L_2) = 25L_1 + 10L_2 \leq 80$$
$$L_1 \geq 0, \quad L_2 \geq 0$$

Com que el problema proposat ens demana maximitzar la funció objectiu en lloc de minimitzar (el que fa `linprog`), hem d'agafar

$$-f_0(L_1, L_1) = -15L_1 - 10L_2$$

i procedir amb la minimització d'aquesta funció objectiu.

Que ens retorna per pantalla:

```
Optimization terminated.

x =

        2.5455
        1.6364

fval =

      -54.5444
```

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 hectàrees (ha) per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

1. Determineu el màxim benefici si les parcel·les d'1 ha generen un guany de 1000 €, les de 4 ha de 2500 €, i les de 6 ha de 3000 €, indicant la configuració de parcel·lació òptima, i si encara queda terreny per parcel·lar o s'aprofiten totes les 100 ha. Teniu en compte que la normativa local d'urbanisme impedeix que una superfície de 100 ha es dividisca en més de 20 parcel·les, i a més obliga que hi haja almenys 5 parcel·les de cada tipus.
2. Què passaria si la normativa canvia i es cancel·la la prohibició que hi haja un màxim de 20 parcel·les?
3. I si partint de la situació descrita a l'apartat anterior s'inclou la prohibició de construir més de 10 parcel·les d'1 ha?
4. Si fóreu el promotor, quina de les dues opcions següents us generaria més beneficis:
    1. Poder parcel·lar fins a 150 ha, però amb les restriccions d'un màxim de 25 parcel·les, un mínim de 5 de cada tipus i no més de 10 parcel·les d'1 ha.
    2. Poder parcel·lar només 100 ha, però sense cap restricció respecte a la quantitat i configuració dels tres tipus de parcel·les possibles.

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

1. Determineu el màxim benefici si les parcel·les d'1 ha generen un guany de 1000 €, les de 4 ha de 2500 €, i les de 6 ha de 3000 €, indicant la configuració de parcel·lació òptima, i si encara queda terreny per parcel·lar o s'aprofiten totes les 100 ha. Teniu en compte que la normativa local d'Urbanisme impedeix que una superfície de 100 ha es dividisca en més de 20 parcel·les, i a més obliga que hi haja almenys 5 parcel·les de cada tipus.

1. Identifiquem les variables del problema (variables de decisió):
   - Nombre de parcel·les d'1 ha: n_1
   - Nombre de parcel·les de 4 ha: n_4
   - Nombre de parcel·les de 6 ha: n_6
2. Identifiquem les restriccions del problema:
   - No podem fer més de 20 parcel·les entre tots els tipus:

     $$f_1(n_1, n_4, n_6) = n_1 + n_4 + n_6 \leq 20$$

   - La suma de les àrees de totes les parcel·les podrà ser, com a molt, 100 ha:

     $$f_2(n_1, n_4, n_6) = 1n_1 + 4n_4 + 6n_6 \geq 100$$

   - Normativa: *Almenys* 5 parcel·les de cada tipus: $n_1 \geq 5$, $n_4 \geq 5$, $n_6 \geq 5$.
3. Plantegem la funció objectiu (beneficis):

   $$f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6$$

Programació lineal: Resolució amb Matlab. Exemple (II)

Plantejament matemàtic del problema:

$$
\begin{aligned}
\text{maximitzar} \quad & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\
\text{subjecte a} \quad & f_1(n_1, n_4, n_6) = n_1 + n_4 + n_6 \leq 20 \\
& f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\
& n_1 \geq 5, \quad n_4 \geq 5, \quad n_6 \geq 5
\end{aligned}
$$

Amb Matlab farem servir la següent sintaxi:

```
>> f = [-1000; -2500; -3000]; Noteu el canvi de signe!
>> A = [1 1 1; 1 4 6];
>> b = [20; 100];
>> lb = [5 5 5];
>> [x, fval] = linprog(f, A, b, [], [], lb, [], [])
```

Programació lineal: Resolució amb Matlab. Exemple (II)

Plantejament matemàtic del problema:

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_1(n_1, n_4, n_6) = n_1 + n_4 + n_6 \leq 20 \\ & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \qquad n_1 \geq 5, \quad n_4 \geq 5, \quad n_6 \geq 5 \end{array}$$

Que ens retorna per pantalla:

```
Optimization terminated.

x =
             5.0000
             5.0000
            10.0000

fval =
        -4.7500e+04
```

El benefici és de 47.500€ quan fem una parcel·lació de 5 parcel·les d'1 ha, 5 parcel·les de 4 ha i 10 parcel·les de 6 ha.

La superfície parcel·lada serà: $1 \cdot n_1 + 4 \cdot n_4 + 6 \cdot n_6 = 85\,ha < 100\,ha$. Per tant, no aprofitem tot el terreny.

Com que $n_1 + n_4 + n_6 = 20$, s'han construït totes les parcel·les possibles.

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

2. Què passaria si la normativa canvia i es cancel·la la prohibició que hi haja un màxim de 20 parcel·les?

Resolem el problema, però sense la restricció $f_1(n_1, n_4, n_6)$.

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1,n_4,n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_2(n_1,n_4,n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \qquad n_1 \geq 5, \quad n_4 \geq 5, \quad n_6 \geq 5 \end{array}$$

Amb Matlab farem servir la següent sintaxi:

```
>> f = [-1000; -2500; -3000]; Noteu el canvi de signe!
>> A = [1 4 6];
>> b = [100];
>> lb = [5 5 5];
>> [x, fval] = linprog(f, A, b, [], [], lb, [], [])
```

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

2 Què passaria si la normativa canvia i es cancel·la la prohibició que hi haja un màxim de 20 parcel·les?

Resolem el problema, però sense la restricció $f_1(n_1, n_4, n_6)$.

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \qquad n_1 \geq 5, \quad n_4 \geq 5, \quad n_6 \geq 5 \end{array}$$

Que ens retorna per pantalla:

```
Optimization terminated.

x =
            50.0000
             5.0000
             5.0000

fval =
        -7.7500e+04
```

El benefici és de 77.500€ quan fem una parcel·lació de 50 parcel·les d'1 ha, 5 parcel·les de 4 ha i 5 parcel·les de 6 ha (construïm 60 parcel·les en total).
La superfície parcel·lada serà: $1 \cdot n_1 + 4 \cdot n_4 + 6 \cdot n_6 = 100\, ha$. Per tant, aprofitem tot el terreny.

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

3. I si partint de la situació descrita a l'apartat anterior s'inclou la prohibició de construir més de 10 parcel·les d'1 ha?

Resolem el problema sense la restricció $f_1(n_1, n_4, n_6)$, i afegint que $n_1 \leq 10$

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \qquad 5 \leq n_1 \leq 10, \quad n_4 \geq 5, \quad n_6 \geq 5 \end{array}$$

Amb Matlab farem servir la següent sintaxi:

```
>> f = [-1000; -2500; -3000]; 
```
Noteu el canvi de signe!

```
>> A = [1 4 6];
>> b = [100];
>> lb = [5 5 5];
>> ub = [10 inf inf];
```

Noteu com s'especifica que 2 de les variables no estan acotades superiorment

```
>> [x, fval] = linprog(f, A, b, [], [], lb, ub, [])
```

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

3. I si partint de la situació descrita a l'apartat anterior s'inclou la prohibició de construir més de 10 parcel·les d'1 ha?

Resolem el problema sense la restricció $f_1(n_1, n_4, n_6)$, i afegint que $n_1 \leq 10$

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \quad 5 \leq n_1 \leq 10, \quad n_4 \geq 5, \quad n_6 \geq 5 \end{array}$$

Que ens retorna per pantalla:

```
Optimization terminated.

x =
            10.0000
            15.0000
             5.0000

fval =
        -6.2500e+04
```

El benefici és de 62.500€ quan fem una parcel·lació de 10 parcel·les d'1 ha, 15 parcel·les de 4 ha i 5 parcel·les de 6 ha (construïm 30 parcel·les en total). Superfície parcel·lada: $1 \cdot n_1 + 4 \cdot n_4 + 6 \cdot n_6 = 100\,ha$. S'aprofita tot el terreny.

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

4. Si fóreu el promotor, quina de les dues opcions següents us generaria més beneficis:
   1. Poder parcel·lar fins a 150 ha però amb les restriccions d'un màxim de 25 parcel·les, un mínim de 5 de cada tipus i no més de 10 parcel·les d'1 ha.

Resolem el problema canviant les restriccions f_1, f_2 i afegint $n_i \leq 10 (i = 1, 4, 6)$

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_1(n_1, n_4, n_6) = n_1 + n_4 + n_6 \leq 25 \\ & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 150 \\ & \qquad 5 \leq n_1 \leq 10, \quad 5 \leq n_4 \leq 10, \quad 5 \leq n_6 \leq 10 \end{array}$$

Amb Matlab farem servir la següent sintaxi:

```
>> f = [-1000; -2500; -3000]; Noteu el canvi de signe!
>> A = [1 1 1; 1 4 6];
>> b = [25; 150];
>> lb = [5 5 5]; ub = [10 10 10];
>> [x, fval] = linprog(f, A, b, [], [], lb, ub, [])
```

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

4. Si fóreu el promotor, quina de les dues opcions següents us generaria més beneficis:
   1. Poder parcel·lar fins a 150 ha però amb les restriccions d'un màxim de 25 parcel·les, un mínim de 5 de cada tipus i no més de 10 parcel·les d'1 ha.

Resolem el problema canviant les restriccions f_1, f_2 i afegint $n_i \leq 10 (i = 1, 4, 6)$

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_1(n_1, n_4, n_6) = n_1 + n_4 + n_6 \leq 25 \\ & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 150 \\ & \qquad 5 \leq n_1 \leq 10, \quad 5 \leq n_4 \leq 10, \quad 5 \leq n_6 \leq 10 \end{array}$$

```
Optimization terminated.

x =
             5.0000
            10.0000
            10.0000

fval =
        -6.0000e+04
```

El benefici és de 60.000€ quan fem una parcel·lació de 5 parcel·les d'1 ha, 10 parcel·les de 4 ha i 10 parcel·les de 6 ha (construïm 25 parcel·les en total).
Superfície parcel·lada: $1 \cdot n_1 + 4 \cdot n_4 + 6 \cdot n_6 = 105\, ha$. S'aprofita tot el terreny.

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

4. Si fóreu el promotor, quina de les dues opcions següents us generaria més beneficis:
    2. Poder parcel·lar només 100 ha però sense cap restricció respecte a la quantitat i configuració dels tres tipus de parcel·les possibles.

Resolem el problema canviant les restriccions f_1, f_2 i afegint $n_i \leq 10 (i = 1, 4, 6)$

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \qquad 0 \leq n_1, \quad 0 \leq n_4, \quad 0 \leq n_6 \end{array}$$

Amb Matlab farem servir la següent sintaxi:

```
>> f = [-1000; -2500; -3000]; Noteu el canvi de signe!
>> A = [1 4 6];
>> b = [100];
>> lb = [0 0 0];
>> [x, fval] = linprog(f, A, b, [], [], lb, [], [])
```

Programació lineal: Resolució amb Matlab. Exemple (II)

Es disposa d'una superfície de 100 ha per a parcel·lar. Es poden fer tres tipus de parcel·les: d'1, 4 i 6 ha.

4 Si fóreu el promotor, quina de les dues opcions següents us generaria més beneficis:

   2 Poder parcel·lar només 100 ha però sense cap restricció respecte a la quantitat i configuració dels tres tipus de parcel·les possibles.

Resolem el problema canviant les restriccions f_1, f_2 i afegint $n_i \leq 10 (i = 1, 4, 6)$

$$\begin{array}{ll} \text{maximitzar} & f_0(n_1, n_4, n_6) = 1000n_1 + 2500n_4 + 3000n_6 \\ \text{subjecte a} & f_2(n_1, n_4, n_6) = n_1 + 4n_4 + 6n_6 \leq 100 \\ & \qquad 0 \leq n_1, \quad 0 \leq n_4, \quad 0 \leq n_6 \end{array}$$

```
Optimization terminated.
x =
        100.0000
          0.0000
          0.0000
fval =
     -1.0000e+05
```

Aquest cas és millor perquè el benefici és de 100.000€ quan fem una parcel·lació en què només hi ha 100 parcel·les d'1 ha.
Superfície parcel·lada: $1 \cdot n_1 + 4 \cdot n_4 + 6 \cdot n_6 = 100\, ha$. S'aprofita tot el terreny.

Programació lineal: Resolució amb Matlab

¶ **Exercici 3:**

$$\begin{array}{ll} \text{minimitzar} & f_0(x) = -5x_1 - 4x_2 - 6x_3 \\ \text{sujecte a} & x_1 - x_2 + x_3 \leq 20 \\ & 3x_1 + 2x_2 + 4x_3 \leq 42 \\ & 3x_1 + 2x_2 \leq 40 \\ & x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0 \end{array}$$

¶ **Exercici 4:** Una dieta saludable ha de contenir m nutrients diferents en quantitats almenys iguals a $l_1, \ldots, l_m$. Podem compondre tal dieta triant quantitats no negatives de n aliments diferents. Una unitat de quantitat de menjar j conté una quantitat a_{ij} de nutrient i, i té un cost c_j.

1. Formuleu el problema lineal que permeta determinar la dieta més barata que satisfaça els requeriments nutricionals esmentats.
2. Formuleu el problema novament suposant que hi ha un cert nutrient k ($k \leq m$) del qual només podem administrar una quantitat exacta en la dieta.
3. Formuleu el problema de nou suposant que les quantitats de cadascun dels nutrients en una dieta equilibrada no poden ser superiors a $u_1, \ldots, u_m$.