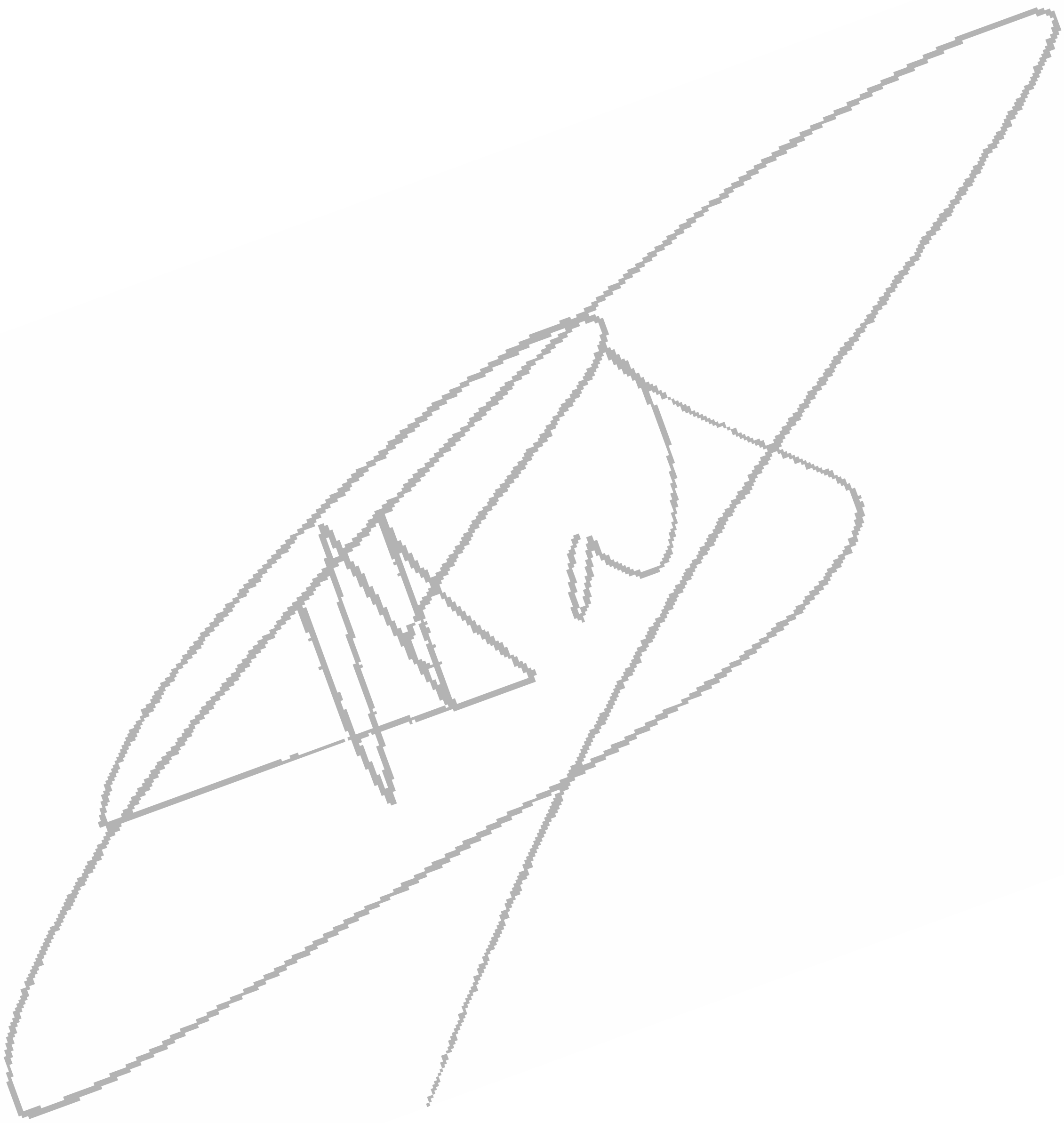
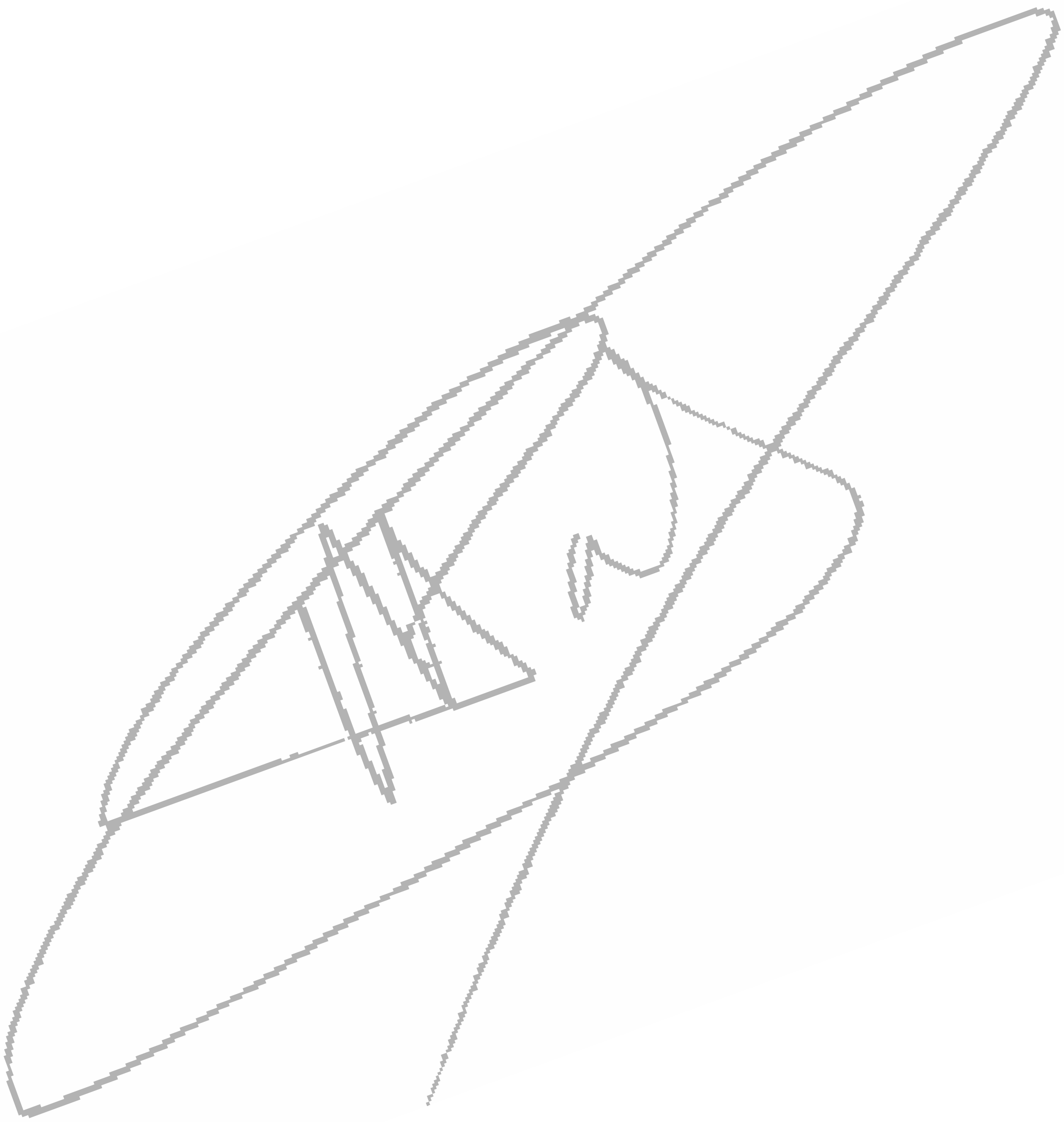
9. Normas y métodos iterativos de resolución de sistemas lineales.

Manuel Palacios
Departamento de Matemática Aplicada
Centro Politécnico Superior
Universidad de Zaragoza

Otoño 2010



Contents



Bibliography

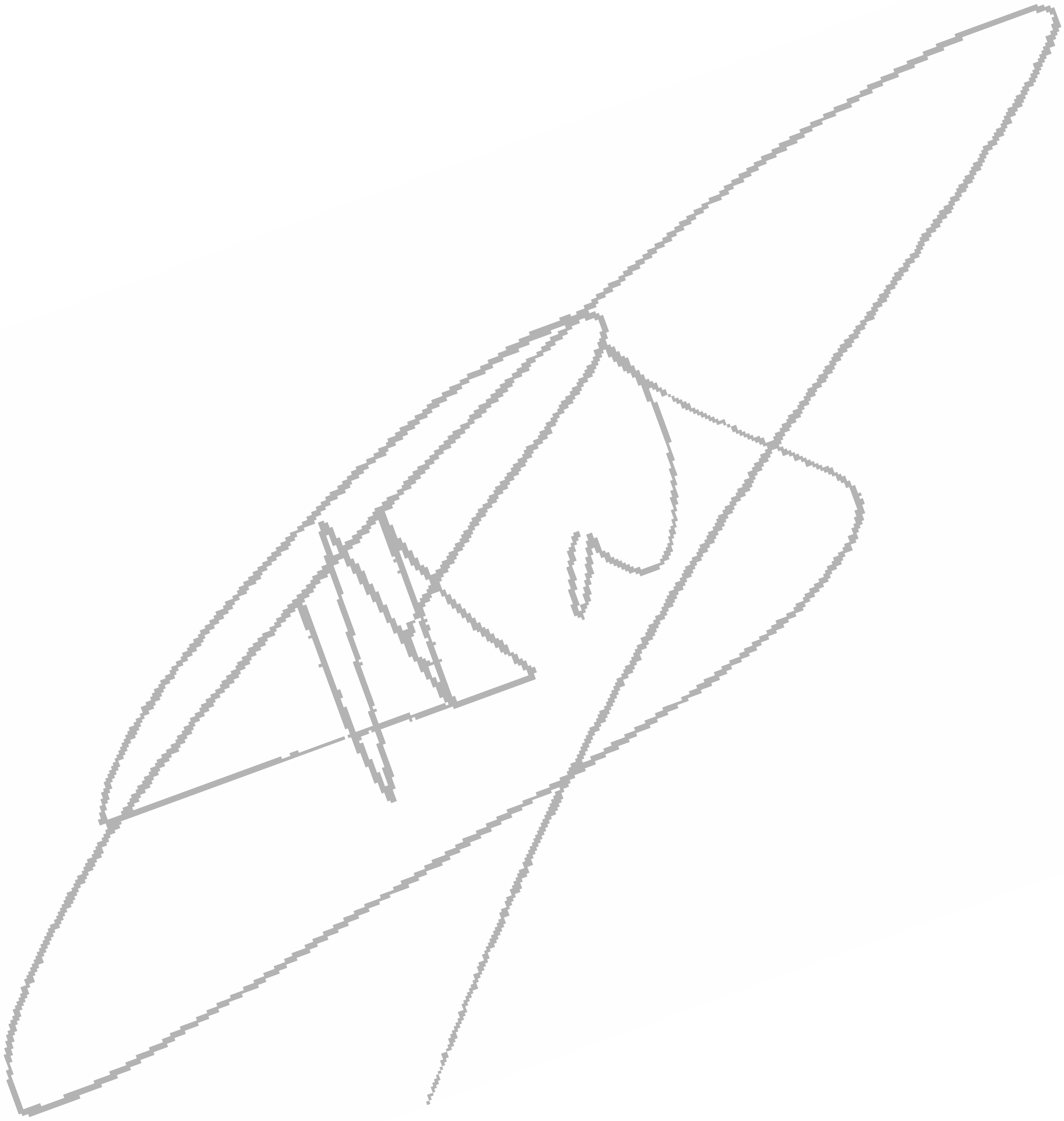
[1] Burden, R. L. and Faires, J. D.: Análisis Numérico. (Fifth ed.) *PWS-Kent Publ. Co.*, 2006.

[2] Palacios, M.: Enunciados de ejercicios de Matemáticas II, grado Ing. Tecn. Industr. http://pcmap.unizar.es/∼mpala. 2010

[3] Villa, A. de la: Problemas de Álgebra. *CLAGSA*. 1988.

[4] Arvesu, J.; Marcellán, F. y Sánchez J.: Problemas resueltos de Álgebra lineal. Paso a paso. *Thomson*. 2005.

[5] Gasca, M.: Cálculo numérico: resolución de ecuaciones y sistemas. *Librería Central*, 1987.



Chapter 9

Normas y métodos iterativos de resolución de sistemas lineales.

En este capítulo introduciremos los conceptos de norma matricial, condicionamiento de una matriz y de convergencia de métodos iterativos para sistemas de ecuaciones lineales.

9.1 Normas vectoriales y matriciales

Recordando las propiedades elementales de la norma eucídea, podemos dar una definición más general de norma, tanto vectorial como matricial (cf. [5] Gasca, [3] de la Villa y [1] Burden-Faires

Definición 9.1.1 *Una* **norma vectorial** *en un espacio vectorial E sobre $\mathbb{R}$ es una aplicación*

$$||\cdot|| : E \longrightarrow \mathbb{R}^+ \cup \{0\}$$

que verifica los axiomas:

1) $||x|| = 0 \Leftrightarrow x = 0$

2) $||t\,x|| = |t|\,||x||$

3) $||x+y|| \leq ||x|| + ||y||$ *(desigualdad triangular)*

Al par $(E, ||\cdot||)$ se le denomina **espacio normado**.

En particular, todo espacio vectorial euclídeo es un espacio normado con la norma euclídea $||x|| = +\sqrt{x \cdot x}$

Ejemplo 9.1.2 *En* $\mathbb{R}^n$,

$$||x||_1 = |x_1| + \ldots + |x_n| = \sum_{i=1}^{n} |x_i|$$

es una norma denominada **norma de la suma** *o taxi-cab.*

Ejemplo 9.1.3 *En* $\mathbb{R}^n$,

$$||x||_1 = (|x_1| + \ldots + |x_n|)^{1/2} = (\sum_{i=1}^{n} |x_i|)^{1/2}$$

es una norma denominada **norma euclídea**.

Ejemplo 9.1.4 *En* $\mathbb{R}^n$,

$$||x||_\infty = \max\{|x_1|, \ldots, |x_n|\}$$

es una norma denominada **norma del máximo** *o de Chebyshev.*

Las bolas de centro el origen y radio unidad son, respectivamente, el cuadrado con vértices en los ejes, el círculo unidad y el cuadrado de lados paralelos a los ejes.

Estas normas definen los mismos conjuntos abiertos y cerrados, es decir, la misma topología. Se dice que son **normas equivalentes**.

Teorema 9.1.5 *(Equivalencia de normas vectoriales)*
Dos normas vectoriales son equivalentes $\Longleftrightarrow$ *existen* $\alpha, \beta \in \mathbb{R}$ *tales que*

$$\alpha\, ||x|| \leq ||x||_\star \leq \beta\, ||x||$$

Las tres normas vectoriales anteriores son equivalentes.

Definición 9.1.6 *Se denomina* **norma matricial** *a una aplicación*

$$||\cdot|| : \mathcal{M}_{\mathrm{R}}(n) \longrightarrow \mathbb{R}^+ \cup \{0\}$$

que verifica los axiomas:
1) $||A|| = 0 \Leftrightarrow A = 0$
2) $||t\, A|| = |t|\, ||A||$
3) $||A + B|| \leq ||A|| + ||B||$ *(desigualdad triangular)*
4) $||A\, B|| \leq ||A||\, ||B||$ *(consistencia)*

Cuando se utilizan normas matriciales y vectoriales conjuntamente es importante que se cumpla también:
5) $||A\, x|| \leq ||A||\, ||x||$ (compatibilidad)

Una forma de construir normas matriciales compatilbes con normas vectoriales es considerar la norma siguiente:

$$||A|| = \sup_{x \neq 0} \frac{||A\, v||}{||x||}$$

Ejercicio 9.1.7 *Comprobar que la anterior es realmente una norma matricial.*

Definición 9.1.8 *La norma anterior se denomina* **norma matricial inducida** *o subordinada por la norma vectorial.*

Inducidas por las normas vectoriales habituales son las siguientes:

Ejemplo 9.1.9 *.*
1) $||A||_1 = \max_{j=1}^n \sum_{i=1}^n |a_{ij}|$ *llamada* **norma de la suma de columnas**
2) $||A||_2 = +\sqrt{\rho(A^T\, A)}$ *llamada* **norma espectral**
3) $||A||_\infty = \max_{i=1}^n \sum_{j=1}^n |a_{ij}|$ *llamada* **norma de Chebyshev**

También se pueden definir normar matriciales no inducidas, por ejemplo, $||A|| = tr\, A$.

Ejercicio 9.1.10 *Encontrar norma 1, espectral y de Chebyshev de las matrices siguientes:*

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & -1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ -1 & 0 & 0 \end{bmatrix}$$

El radio espectral de una matriz, $\rho(A)$ no define ninguna norma matricial, ya que, por ejemplo, con las matrices A $= \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$ y B $= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ no se cumple la desigualdad triangular. Sin embargo, está muy cerca de serlo.

Teorema 9.1.11 *(cf. [1]Burden, [3]de la Villa)*
1) $\rho(A) < ||A||$, para cualquier norma,
2) $\forall \epsilon > 0$, existe una norma inducida tal que $||A|| \leq \rho(A) + \epsilon$

Teorema 9.1.12 *(de Banach) cf. [5]Gasca*
1) Si $I + B$ es singular $\Longrightarrow ||B|| \geq 1$
2) Si $||B|| < 1 \Longrightarrow I + B$ regular y

$$\frac{1}{1 + ||B||} \leq ||(I + B)^{-1}|| \leq \frac{1}{1 - ||B||}$$

Teorema 9.1.13 *Son equivalentes las siguientes afirmaciones:*
1) $\rho(A) < 1$
2) $A^n \longrightarrow 0$
3) $||A||^n \longrightarrow 0????$
4) $||A^n v|| \longrightarrow 0, \quad \forall v$

9.1.1 Ejercicios propuestos

Ejercicios de Palacios [2]

9.2 Condicionamiento

9.2.1 Condición de un problema

Considérese una aplicación $\mathcal{P} : \mathbb{R}^n \longrightarrow \mathbb{R}$, es decir, el problema de calcular $\mathcal{P}(x)$ para los datos $x = (x_1, \ldots, x_n)$. ¿Cómo influyen las perturbaciones en los datos x sobre el resultado $\mathcal{P}(x)$?.

Definición 9.2.1 *Se denomina condición, κ, de un problema $\mathcal{P}$ al número real más pequeño tal que*

$$\frac{|\hat{x}_i - x_i|}{|x_i|} \leq eps \Longrightarrow \frac{|\mathcal{P}(\hat{x}) - \mathcal{P}(x)|}{|\mathcal{P}(x)|} \leq \kappa \cdot eps \tag{9.1}$$

Se dice que un problema $\mathcal{P}$ está bien condicionado, si κ no es muy grande, en caso contrario, $\mathcal{P}$ está mal condicionado.

En esta definición, *eps* representa un número pequeño; si *eps* fuese la precisión del ordenador (por ejemplo, 10^{-6} en simple precisión), $\hat{x}_i$ puede ser considerado como el número de máquina de x_i. Téngase en cuenta que el número de condición no depende del algoritmo utilizado para resolver el problema $\mathcal{P}$, sino que depende de los datos x_i y del propio problema $\mathcal{P}$.

Ejemplo 9.2.2 *Multiplicación de dos números reales. Dados* $x_i, i = 1, 2$, *considérese el problema de calcular* $\mathcal{P}(x_1, x_2) = x_1 \cdot x_2$.

Tomando los valores perturbados:

$$\hat{x}_1 = x_1(1 + \epsilon_1), \quad \hat{x}_2 = x_2(1 + \epsilon_2), \quad |\epsilon_i| \leq eps$$

se tiene

$$\frac{\hat{x}_1 \cdot \hat{x}_2 - x_1 \cdot x_2}{x_1 \cdot x_2} = (1 + \epsilon_1) \cdot (1 + \epsilon_2) - 1 = \epsilon_1 + \epsilon_2 + \epsilon_1 \cdot \epsilon_2$$

Como *eps* es un número pequeño, $\epsilon_1 \cdot \epsilon_2$ se puede despreciar frente $|\epsilon_1| + |\epsilon_2|$; por lo tanto,

$$|\frac{\hat{x}_1 \cdot \hat{x}_2 - x_1 \cdot x_2}{x_1 \cdot x_2}| \leq 2 \cdot eps$$

resulta, $\kappa = 2$, es decir, un problema bien condicionado.

Ejemplo 9.2.3 *Substracción de dos números reales. Dados* $x_i, i = 1, 2$, *considérese el problema de calcular* $\mathcal{P}(x_1, x_2) = x_1 - x_2$.

En forma análoga al anterior se obtiene:

$$|\frac{(\hat{x}_1 - \hat{x}_2) - (x_1 - x_2)}{x_1 - x_2}| = |\frac{x_1 \cdot \epsilon_1 - x_2 \cdot \epsilon_2}{x_1 - x_2}| \leq \frac{|x_1| + |x_2|}{|x_1 - x_2|} \cdot eps$$

En consecuencia, si sign $x_1 = -$ sign x_2, lo que corresponde a una suma, no a una resta, se tiene $\kappa = 1$, es decir, un problema bien condicionado.

Por el contrario, si $x_1 \approx x_2$, el número de condición resulta muy grande y el problema muy mal condicionado. Por ejemplo,

$$x_1 = \frac{1}{51}, \quad x_2 = \frac{1}{52} \Longrightarrow \kappa \approx \frac{2/50}{(1/50)^2} = 100$$

Si en esta situación, se realizan las operaciones con 3 cifras exactas (en base 10), se obtiene: $\hat{x}_1 = 0.196 \cdot 10^{-1}$, $\hat{x}_2 = 0.192 \cdot 10^{-1}$ y $\hat{x}_1 - \hat{x}_2 = 0.400 \cdot 10^{-3}$. Como las dos primeras cifras de $\hat{x}_1$ y $\hat{x}_2$ son las misma, la substracción las hace desaparecer y solo se conserva una cifra significativa. Se dice que ha habido *pérdida de cifras significativas*.

9.2.2 Condicionamiento de un sistema lineal

Sabemos que el condicionamiento influye en la calidad de la solución de un problema cualquiera. En particular, en el problema de hallar la solución de un sistema lineal nos encontramos con que al comparar el valor exacto del término independiente de un sistema con el calculado puede haber discrepancias. En concreto, definiendo el vector residual r en la forma

$$r = \tilde{b} - b,$$

en donde $\tilde{b}$ es el valor calculado, resulta:

Teorema 9.2.4 *Si A es una matriz regular, se verifica:*
1) $\|\tilde{x} - x\| \leq \|r\| \, \|A^{-1}\|$

2) $\dfrac{\|\tilde{x} - x\|}{\|x\|} \leq \; A\| \, \|A^{-1}\| \, \dfrac{\|r\|}{\|b\|}$

Dem.: Como $r = \tilde{b} - b = A\,\tilde{x} - A\,x = A\,(\tilde{x} - x)$, resulta

$$\|\tilde{x} - x\| = \|A^{-1} r\| \leq \|r\| \, \|A^{-1}\|$$

Por otro lado, como $A\,x = b$, $\|x\| \geq \dfrac{\|b\|}{\|A\|}$, luego de la desigualdad anterior resulta

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq \frac{\|r\| \, \|A^{-1}\|}{\|x\|} \leq \frac{\|A^{-1}\| \, \|A\| \, \|r\|}{\|b\|}$$

De acuerdo con la definición 9.2.1, podemos dar la siguiente

Definición 9.2.5 *Se denomina número de condición de una matriz al número*

$$\kappa(A) = \|A\| \, \|A^{-1}\|$$

Propiedad 9.2.6 *1)* $\kappa(A) \geq 1$
2) $\kappa(\lambda\, A) = \kappa(A)$
3) $\kappa(A^{-1}) = \kappa(A)$
4) Si Q es ortogonal, $\kappa(Q) = 1$

Si $\kappa(A)$ es pequeño, se dice que la matriz A está bien condicionada, si es grande que A está mal condicionada.

Para encontrar una **estimación** del número de condición,
se resuelve el sistema $A\,x = b$ por eliminación gaussiana utilizando aritmética de p dígitos, obteniendo la solución $\tilde{x}$;
a continuación, se resuelve también con p dígitos el sistema $A\,y = r$ (que es inmediato, ya que se tienen en memoria las operaciones de la eliminación anterior) obteniendo el vector $\tilde{y}$;
finalmente, se usa la desigualdad de Forsythe y Moler siguiente

$$\|r\| \approx 10^{-p} \, \|A\| \, \|\tilde{x}\|$$

para estimar el número de condición por la siguiente

$$\kappa(A) \approx \frac{\|\tilde{y}\|}{\|\tilde{x}\|} \, 10^{p}$$

Este procedimiento proporciona buenas estimaciones del número de condición (cf. Burden y Faires [1]).

En el siguiente teorema veremos la relación entre las perturbaciones de los elementos del sistema y el número de condición de la matriz de coeficientes.

Suponiendo que la matriz A del sistema

$$A\,x = b$$

es regular, el sistema perturbado se puede escribir en la forma

$$(A + \Delta A)\,(x + \Delta x) = b + \Delta b \tag{9.2}$$

Propiedad 9.2.7 *Si*

$$\|\Delta A\|\,\|A^{-1}\| < 1,$$

la matriz de coeficientes $A + \Delta A$ del sistema perturbado es regular

Dem.:

$$A + \Delta A = A\,(I + A^{-1}\,\Delta A)$$

Como A es regular, existe A^{-1}. La matriz $I + A^{-1}\,\Delta A$, como sabemos, será regular si $\|A^{-1}\,\Delta A\| < 1$, para lo cual es suficiente que $\|\Delta A\|\,\|A^{-1}\| < 1$. ■

Teorema 9.2.8 *Se verifica la siguiente acotación:*

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{K(A)}{1 - K(A)\dfrac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right)$$

Corolario 9.2.9 *Si $\|\Delta A\| = 0$, entonces*

$$\frac{\|\Delta x\|}{\|x\|} \leq K(A)\,\frac{\|\Delta b\|}{\|b\|}$$

En consecuencia, errores pequeños en b pueden producir resultados nefastos en la solución x.

Refinamiento iterativo o corrección residual

En el caso de que la matriz de coeficientes esté mal condicionada se puede utilizar el siguiente algoritmo de refinamiento iterativo o corrección residual para mejorar el resultado. Téngase cuidado porque si el problema no está mal condiconado este procedimiento encarece el coste computacional y no mejora sensiblemente la solución.

Algoritmo.

- Encontrar la solución $\tilde{x}$ del sistema $A\,x = b$
- Encontrar el vector residual $r = A\,\tilde{x} - b$ con doble precisón
- Encontrar la solución $\tilde{e}$ de la ecuación $A\,e = r$
- Construir la nueva aproximación de la solución $x = \tilde{x} + \tilde{e}$

9.2.3 Ejercicios propuestos

.

Ejercicios de Palacios [2]
Ejercicios de [4](Arvesu) números 14.7, 14.8, 14.11, 14.20, 14.22
Ejercicios de [1](Burden) números 7.4, 7.5, 7.6

9.3 Métodos iterativos lineales

Consideraremos un sistema lineal

$$A\,x = b, \tag{9.3}$$

siendo A una matriz inversible, es decir, el sistema es compatible determinado.

Como ya se indicó en un capítulo anterior, un método iterativo para la resolución de sistemas lineales permite construir una sucesión $\{x^{(m)}\}_{m=0}^{\infty}$ de vectores que "aproximan" la solución exacta $x = A^{-1}\,b$ del sistema en el sentido de que

$$\lim_{m->\infty} x^{(m)} = x$$

Los métodos iterativos más usados son los que construyen la solución en la forma

$$x^{(m+1)} = B\,x^{(m)} + c, \quad B \in M_{\mathbb{R}}(n), \quad c \in \mathbb{R}^n, \tag{9.4}$$

a partir de $x^{(0)} \in \mathbb{R}^n$ arbitrario.

Definición 9.3.1 *Un método (9.4) se dice convergente, si para cualquier $x^{(0)} \in \mathbb{R}^n$, se verifica:*

$$\lim_{m->\infty} x^{(m)} = x = A^{-1}\,b$$

Una condición mínima que se ha de cumplir es que el método ha de ser consistente, es decir, que, en el límite, la ecuación (9.4) se transforme en

$$x = B\,x + c, \tag{9.5}$$

siendo $x = A^{-1}\,b$ la solución exacta

Teorema 9.3.2 *Un método (9.4) es convergente para el sistema (9.3) si y solo se verifican las dos condiciones siguientes:*
i) el método es consistente con el sistema
ii) $\rho(B) < 1$

Dem.: $\boxed{\Rightarrow}$ Si el método es convergente, pasando al límite en (9.4) resulta la consistencia. Por otro lado, de (9.4) y (9.5), resulta:

$$x^{(m+1)} - x = B\,(x^{(m)} - x) = B^2\,(x^{(m-1)} - x) = ... = B^{m+1}\,(x^{(0)} - x)$$

Como esta diferencia debe tender a 0 por la convergencia, se debe cumplir

$$\lim_{m->\infty} B^m\,v = 0, \quad \forall v \in \mathbb{R}^n,$$

lo cual se cumple si $\rho(B) < 1$ (cf. 9.1.13).

$\boxed{\Leftarrow}$ Si se cumplen ii), debe existir alguna norma tal que $\| B \| < 1$. Consideremos una norma vectorial compatible con ella. Como más arriba, se tiene:

$$x^{(m)} - x = B^m \, (x^{(0)} - x)$$

Tomando normas resulta:

$$\|x^{(m)} - x\| = \|B\|^m \, \|x^{(0)} - x\|$$

Pasando al límite, se tiene la convergencia. ■

9.3.1 Estabilidad

Estudiar la estabilidad de un método iterativo es analizar la influencia de perturbaciones (pequeñas modificaciones). Así que, supongamos que los cálculos se realizan cometiendo un pequeño error, es decir, en vez de

$$x^{(m)} = B \, x^{(m-1)} + c,$$

se calcula

$$\tilde{x}^{(m)} = B \, \tilde{x}^{(m-1)} + c + \rho^{(m)}$$

En consecuencia, se tendrá

$$\begin{aligned} \tilde{x}^{(m)} - x^{(m)} &= B \, (\tilde{x}^{(m-1)} - x^{(m)}) + \rho^{(m)} \\ \tilde{x}^{(0)} - x^{(0)} &= \rho^{(0)}; \end{aligned}$$

de donde

$$\tilde{x}^{(m)} - x^{(m)} = B^m \, \rho^{(0)} + \sum_{j=1}^{m} B^{m-j} \, \rho^{(j)},$$

y tomando normas compatibles:

$$\begin{aligned} \|\tilde{x}^{(m)} - x^{(m)}\| &\le \textstyle\sum_{j=1}^{m} \|B\|^{m-j} \, \|\rho^{(j)}\| \le sup\{\|\rho^{(j)}\|\} \textstyle\sum_{j=1}^{m} \|B\|^j \le \\ &\le sup\{\|\rho^{(j)}\|\} \textstyle\sum_{j=1}^{\infty} \|B\|^j = \dfrac{1}{1 - \|B\|} \, sup\{\|\rho^{(j)}\|\} \end{aligned}$$

Por lo tanto, pequeños errores de redondeo producirán errores pequeños en el resultado en la iteración m y el método será estable, siempre que sea convergente.

Notemos que un valor de $\|B\| < 1$, pero próximo a 1 producirá una cota muy grande. Interesa que $\|B\|$ sea pequeña.

9.3.2 Velocidad de convergencia

Más arriba hemos obtenido

$$\|x^{(m)} - x\| = \|B^m\| \, \|x^{(0)} - x\|$$

Esto puede ser interpretado como que en cada una de las m primeras iteraciones el error se ha reducido en un factor de $\|B^m\|^{1/m}$ y, en consecuencia, se puede estimar que para que el error se reduzca en un factor de 10^{-t} se deben realizar N iteraciones, cumpliéndose:

$$(\|B^m\|^{1/m})^N \le 10^{-t}, \quad \text{o bien} \quad N \ge \frac{t}{-log_{10}(\|B^m\|^{1/m})}$$

Definición 9.3.3 *El número*

$$R(B^m) = -\log_{10} \|B^m\|^{1/m} \tag{9.6}$$

es denominado **velocidad media de convergencia** *en m iteraciones.*

Se puede demostrar que

$$\rho(B) = \lim_{m->\infty} \|B^m\|^{1/m}$$

Definición 9.3.4 *Al número* $-\log_{10}(1/\rho(B))$ *se le denomina* **velocidad asintótica** *de convergencia.*

Ejercicio 9.3.5 *Encontrar la velocidad de convergencia en 5 iteraciones del método iterativo que toma $M = D$ (diagonal principal de A) al aplicarlo a la resolución del sistema cuya matriz ampliada es la siguiente:*

$$A = \begin{pmatrix} 800.2669 & 0.8648 & 0.8642 \\ 0.2161 & 800.1441 & 0.1440 \end{pmatrix}$$

Solución: La matriz de iteración B y su potencia serán:

$$B = M^{-1}\,N = \begin{pmatrix} 0 & -0.0010906 \\ -0.000270076 & 0 \end{pmatrix}, \quad B^5 = \begin{pmatrix} 0 & -9.203 * 10^{-17} \\ -2.30032 * 10^{-17} & 0 \end{pmatrix}$$

Así que

$$\|B^5\|_\infty = 9.203 * 10^{-17}$$

Y por lo tanto,

$$R(B^5) = 3.20721$$

El valor propio de B de módulo mayor es $\lambda = 0.000540226$, por lo que la velocidad asintótica será

$$\log_{10}(1/\rho(B)) = 3.26742$$

Como se puede observar ambas son muy cercanas.■

9.3.3 Cota del error

Se puede demostrar el siguiente resultado
Teorema 9.3.6

$$\|x^{(m)} - x\| \leq \frac{\|B\|}{1 - \|B\|} \|x^{(m)} - x^{(m-1)}\|$$

Dem.: Tomando normas en la siguiente expresión

$$x^{(m)} - x = x^{(m)} - x^{(m+1)} + x^{(m+1)} - x,$$

resulta

$$\|x^{(m)} - x\| \leq \|x^{(m)} - x^{(m+1)}\| + \|x^{(m+1)} - x\| \leq \|x^{(m+1)} - x^{(m)}\| + \|B\|\,\|x^{(m)} - x\|,$$

es decir,

$$(1 - \|B\|)\,\|x^{(m)} - x\|\| \leq \|x^{(m+1)} - x^{(m)}\| \leq \|B\|\,\|x^{(m)} - x^{(m-1)}\|,$$

de donde se deduce la desigualdad buscada.■

Esta acotación suele utilizarse como test de parada del algoritmo.

9.4 Resultados de convergencia para los métodos iterativos usuales

Los métodos iterativos más usuales son los métodos de Jacobi, Gauss-Seidel y relajación. Todos ellos fueron introducidos en un capitulo precedente. Para su definición, descomponemos la matriz A en la forma

$$A = D - L - U,$$

donde D es la matriz diagonal principal de A, $-L$ es una matriz triangular inferior estricta cuyos elementos no nulos coinciden con los de A y $-U$ es una matriz triangular superior estricta cuyos elementos no nulos coinciden con los de A.

El método de **Jacobi** se define considerando la matriz $M = D$.

El método de **Gauss-Seidel** se define considerando la matriz $M = D - L$

Recordemos que para determinar la aproximación $x^{(m+1)}$, en el método de Jacobi, se utilizan los $x_i^{(m)}$ de la iteración anterior, mientras que, en el método de Gauss-Seidel, se utilizan los últimos $x^{(m+1)}$ calculados. Por eso, en el primero se requiere mantener en memoria dos vectores, el $x^{(m+1)}$ y el $x^{(m)}$, mientras que en el de Gauss-Seidel solo se necesita uno de ellos.

Ejemplo 9.4.1 *Resolver mediante los métodos de Jacobi y Gauss-Seidel el sistema $Ax = b$ siguiente:*

$$\bar{A} = \begin{pmatrix} 10 & -1 & 2 & 0 & 6 \\ -1 & 11 & -1 & 3 & 25 \\ 2 & -1 & 10 & -1 & -11 \\ 0 & 3 & -1 & 8 & 15 \end{pmatrix}$$

Ayuda:

En este ejemplo, el método de Gauss-Seidel converge más rapidamente que el de Jacobi, aunque esto no es así en general. ■

9.4.1 Matrices estrictamente diagonal dominantes

Definición 9.4.2 *Una matriz A se dice* **diagonal dominante** *si se cumple:*

$$|a_{kk}| \leq \sum_{j=1, j\neq k}^{n} |a_{kj}|, \quad \forall k = 1, 2, ..., n$$

Se dice **diagonal dominante estricta** *si se cumple la desigualdad estrictamente.*

Teorema 9.4.3 *Para esta matrices, se cumple que los métodos de Gauss-Seidel y de Jacobi son convergentes y $\rho(J) \leq K$, y $\rho(G) \leq K$, siendo*

$$K = \max\{\sum_{j=1, j\neq i}^{n} \frac{|a_{ij}|}{|a_{ii}|}\} < 1$$

Dem.: Consultar [1]. ■

En el ejemplo (9.4.1) se tiene:

$$\rho(J) = 0.4264, \quad \rho(G) = 0.0898 \text{ y } K = 0.5$$

9.4.2 Matrices simétricas definidas positivas

Teorema 9.4.4 *Si A es simétrica y $a_{ii} > 0, \quad \forall i$, entonces, el método de Gauss-Seidel converge $\Longleftrightarrow$ A es definida positiva*

Dem.: Consultar [1]. ■

Teorema 9.4.5 *Si A es simétrica y $D + L + U$ es definida positiva, entonces, el método de Jacobi converge $\Longleftrightarrow$ A es definida positiva*

Dem.: Consultar [1]. ■

9.4.3 Matrices tridiagonales

Teorema 9.4.6 *En este caso, los radios espectrales de las matrices de iteración de los métodos de Jacobi y Gauss-Seidel verifican: $\rho(G) = \rho(J)^2$ y, por tanto, ambos convergen o ambos divergen. En el caso convergente, el de Gauss-Seidel es, asintóticamente, más rápido que el de Jacobi.*

Dem.: Consultar [1]. ■

9.4.4 Convergencia de los métodos de relajación

Teorema 9.4.7 *(Kahan) Si $a_{ii} \neq 0, \forall i, \quad \rho(G_\omega) \geq |\omega - 1|$*

Dem.: Basta observar que, por construcción de G_w:

$$\det G_\omega = \det D^{-1} (1 - \omega)^n \det D = (1 - \omega)^n = \prod \lambda_i \leq \rho(G_\omega)^n \Rightarrow \rho(G_\omega) \geq |\omega - 1|.$$

Teorema 9.4.8 *El método SOR solo puede converger si $0 < \omega < 2$*

Dem.: Es consecuencia inmediata del teorema anterior. ■

Teorema 9.4.9 *Si A es estrictamente diagonal dominante y $0 < \omega \leq 1 \Longrightarrow$ SOR converge.*

Dem.: Basta calcular la norma de Chebyshev del vector $y = G\,u$, con $\|u\| = 1$. Teniendo en cuenta la definición del método SOR, resulta

$$\|y\|_\infty = \|G\,u\|_\infty \leq \omega\,\|G\|_\infty + (1 - \omega) < 1, \text{ si } 0 < \omega \leq 1,$$

luego

$$\|G\|_\infty = \max_{\|u\|_\infty = 1} \|G\,u\|_\infty < 1. \blacksquare$$

Teorema 9.4.10 *(Ostrowski-Reich)*

Si A es simétrica, real, con $a_{ii} > 0, \forall i$, entonces, SOR converge $\Longleftrightarrow$ A es definida positiva y $0 < \omega < 2$

Teorema 9.4.11 *Si A es tridiagonal simétrica definida positiva, entonces: los métodos de Jacobi, Gauss-Seidel y SOR convergen y el valor óptimo del parámetro ω es*

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(B_J)^2}}$$

Ejercicio 9.4.12 *Estudiar el condicionamiento del sistema, la convergencia y el número de iteraciones necesarias en su caso de los métodos usuales aplicados a la resolución de los sistemas lineales definidos por su matriz ampliada siguiente:*

$$A = \begin{bmatrix} 7 & 3 & 1 & 5 \\ 2 & 1 & -7 & -6 \\ 1 & 4 & 0 & -3 \end{bmatrix}, \; B = \begin{bmatrix} 1 & .4 & .3 & .2 \\ .8 & 1 & .1 & 1.2 \\ .5 & .5 & 1 & 0.3 \end{bmatrix}$$

Tolerancia pedida, Tol = 10^{-3}

Ejercicio 9.4.13 *Estudiar el condicionamiento del sistema, la convergencia y el número de iteraciones necesarias en su caso de los métodos usuales aplicados a la resolución de los sistemas lineales definidos por su matriz ampliada siguiente:*

$$C = \begin{bmatrix} -1 & 1 & 2.1 & 0 \\ 7.2 & 0 & 2 & 7.2 \\ -2.2 & 4 & 1 & 1.8 \end{bmatrix}, \; D = \begin{bmatrix} 1 & .363 & .273 & .18182 \\ .75 & 1 & .125 & 1.25 \\ .5 & .5 & 1 & .25 \end{bmatrix}$$

Tolerancia pedida, Tol = 10^{-3}

9.4.5 Ejercicios propuestos

Ejercicios de [4](Arvesu) números 15.4, 15.6, 15.8, 15.9 y propuestos 15.1, 15.2, 15.3, 15.4, 15.5, 15.6, 15.7 y 15.9

Ejercicios de Palacios [2]