

# Resumen de definiciones y condiciones útiles de numérico.

Garchorin!

November 8, 2018

## 1 Búsqueda de raíces - (Práctica 3)

Notaremos  $x^*$  a el resultado buscado (en este caso las raíces)

### 1.1 Método de la bisección

Sea  $f$  una función continua en el intervalo  $[a_k, b_k]$ , donde  $f(a_k)f(b_k) < 0$  (es decir, tienen distinto signo), entonces  $f$  tiene una raíz en el intervalo  $(a_k, b_k)$ . La idea es agarrar el punto medio  $m_k = (a_k + b_k)/2$  y verificar 3 cosas:

- Si  $m_k \approx x^*$  dependiendo el error que queremos terminamos o no, sino
- $x^* \in (a_k, m_k)$
- $x^* \in (m_k, b_k)$

Si se dan alguna de las últimas dos opciones, se vuelve a iterar utilizando el nuevo intervalo.

### 1.2 Método de la secante

Sea  $f$  una función a la que le queremos encontrar las raíces, dados dos puntos iniciales  $x_0$  y  $x_1$ , trazamos una secante en los puntos formados por  $(x_0, f(x_0))$ ,  $(x_1, f(x_1))$ , para obtener un nuevo punto  $x_2$ , el cual es determinado por la intersección del eje  $x$  con la recta secante recién trazada. En la siguiente iteración, se trazará una secante entre los puntos  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ .

### 1.3 Método del punto fijo

Los métodos vistos se aplican a la solución de la ecuación  $f(x) = 0$ . El método de punto fijo sirve para resolver la ecuación  $g(x) = x$ . Se busca un  $x^*$  tal que su imagen, por medio de la función  $g$ , sea el mismo  $x^*$ . Por tal motivo se dice que  $x^*$  es un punto fijo de la función  $g$ . La aplicación del método es muy sencilla. A partir de un  $x_0$  dado, se aplica varias veces la fórmula:

$$x_{k+1} = g(x_k)$$

Se espera que la sucesión  $\{x_k\}$  construida mediante las iteraciones converja hacia  $x^*$ .

Tenemos  $f(x)$  a la cual queremos buscarle la raíz. En vez de eso, igualaremos  $f(x) = 0$ , y de lo que queda la forma explícita de  $f(x)$ , despejamos  $x$ . Esa nueva función será  $g(x)$ . A esta nueva función, le queremos buscar el punto fijo, es decir,  $g(x) = x$ . Notemos que  $g(x) - x = 0 = f(x)$ . Ahora tenemos que hacerle la derivada a  $g(x)$  y buscar que  $|g'(x)| < 1$ . De ahí, resulta un intervalo donde estará el punto fijo de  $g(x)$ . Ahora a  $x_0$  (una de las puntas del intervalo) le aplicaremos el método en sí, que es  $x_i = g(x_{i-1})$ . Después de  $n$  iteraciones deberíamos llegar al punto fijo de  $g$ , que será la raíz de  $f$ .

### 1.3.1 Teoremas de convergencia del método de punto fijo

**Teorema 1:** Sea  $g$  continuamente diferenciable en el intervalo  $[a, b]$  tal que

$$g([a, b]) \subseteq [a, b] \quad (\forall x \in [a, b] \quad g(x) \in [a, b])$$
$$|g'(x)| < 1 \quad \forall x \in [a, b]$$

Entonces, existe un único  $x^*$  en  $[a, b]$  solución de  $x = g(x)$  y la iteración de punto fijo converge a  $x^*$  para todo  $x_0 \in [a, b]$ .

**Teorema 2:** Sea  $x^*$  solución de  $x = g(x)$ ,  $g$  continuamente diferenciable en un intervalo abierto  $I$  tal que  $x^* \in I$ ,  $|g'(x^*)| < 1$ . Entonces, la iteración de punto fijo converge a  $x^*$  para todo  $x_0$  suficientemente cerca de  $x^*$ .

## 1.4 Método de Newton (Newton-Raphson)

Dado  $x_0$ , se construye la recta tangente en  $(x_0, f(x_0))$ . El valor de  $x$  donde esta recta corta el eje  $x$  es el nuevo valor  $x_1$ . Ahora se construye la recta tangente en el punto  $(x_1, f(x_1))$ . El punto de corte entre la recta y el eje  $x$  determina  $x_2$ , así sucesivamente...

En el caso general, dado  $x_k$ , se construye la recta tangente en el punto  $(x_k, f(x_k))$ :

$$y = f'(x_k)(x - x_k) + f(x_k)$$

Para  $y = 0$ , se tiene  $x = x_{k+1}$  (decimos esto porque en la siguiente iteración buscamos que  $y = 0$ ):

$$0 = f'(x_k)(x_{k+1} - x_k) + f(x_k) \rightarrow x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

## 2 Sistemas de ecuaciones no lineales

### 2.1 Definiciones

- Matriz jacobiana

La matriz jacobiana de la función  $F : R^n \rightarrow R^n$ , denotada por  $JF(x)$  o por  $F'(x)$ , es una matriz de tamaño  $n \times n$ , en la que en la  $i$ -ésima fila están las  $n$  derivadas parciales de  $F_i$

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \dots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \dots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$$

### 2.2 Método de Newton en $R^n$

Un sistema de  $n$  ecuaciones con  $n$  incógnitas se puede escribir de la forma:

$$F_1(x_1, x_2, \dots, x_n) = 0$$

$$F_2(x_1, x_2, \dots, x_n) = 0$$

$$\vdots$$

$$F_n(x_1, x_2, \dots, x_n) = 0$$

donde cada  $F_i$  es una función de  $n$  variables con valor real, o sea,  $F_i : R^n \rightarrow R$ . Denotemos  $x = (x_1, x_2, \dots, x_n)$  y

$$F(x) = \begin{bmatrix} F_1(x) \\ F_2(x) \\ \vdots \\ F_n(x) \end{bmatrix}$$

La idea es igual a la del método de newton en una variable (utilizamos a  $F(X)$  de la misma forma, teniendo como  $F'(x)$  al Jacobiano), para verlo mejor hacemos un paralelismo de como se van deduciendo las fórmulas:

#### Newton en $\mathbb{R}$

$$y = f'(x_k)(x - x_k) + f(x_k)$$

Para  $y = 0$

$$0 = f'(x_k)(x_{k+1} - x_k) + f(x_k) \rightarrow x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

#### Newton en $\mathbb{R}^n$

$$y = F(x_k) + F'(x_k)(x - x_k)$$

$$0 = F(x_k) + F'(x_k)(x_{k+1} - x_k) \rightarrow x_{k+1} = x_k - F'(x_k)^{-1}F(x_k)$$

## 3 Sistemas de ecuaciones lineales

### 3.1 Definiciones

#### 3.1.1 Pivoteo parcial

- Se debe construir una matriz de coeficientes y el vector con los términos independientes, correspondientes al sistema, y se crea una matriz llamada la matriz aumentada.
- Se busca el número mayor (en valor absoluto) en cada la columna correspondiente a la etapa y se procede a un cambio de filas para ubicar el mayor elegido en la posición correspondiente a la etapa.
- Una vez ubicado el número mayor, se procede al cálculo de los multiplicadores correspondientes a la etapa.

Ejemplo: Tenemos la siguiente matriz, y vamos por el segundo paso de la reducción gaussiana:

$$A = \left[ \begin{array}{ccccc|c} 1 & 7 & 9 & -4 & 3 & 5 \\ 0 & \mathbf{-5} & 6 & -4 & -8 & -3 \\ 0 & \mathbf{6} & -3 & 5 & 2 & -7 \\ 0 & \mathbf{-7} & 7 & -2 & 7 & 6 \\ 0 & \mathbf{-2} & 0 & -6 & 3 & -5 \end{array} \right]$$

Los candidatos corresponden a la columna 2 (marcados en negrita):

$$|a_{2,2}| = 5 \quad |a_{2,3}| = 6 \quad |a_{2,4}| = 7 \quad |a_{2,5}| = 2$$

Como el máximo valor corresponde a  $|a_{2,4}| = 7$  que está en la fila 4, se procede a cambiar la fila 2 (la que estamos en el algoritmo en este momento) por la 4 y se prosigue de forma normal.

#### 3.1.2 Autovalores

Sea  $A \in \mathbb{R}^{n \times n}$  una matriz cuadrada real. Un número  $\lambda$ , real o complejo, es un autovalor de  $A$  si existe un vector columna real o complejo no nulo  $v \in \mathbb{C}^{n \times 1}$  tal que  $Av = \lambda v$ . En ese caso, se dice que  $v$  es un autovector asociado al autovalor  $\lambda$ .

Se obtienen calculando las raíces del polinomio característico de  $A$ , donde:

$$p(\lambda) = \det(A - \lambda I)$$

#### 3.1.3 Matriz Hermitiana

Una matriz Hermitiana es una matriz cuadrada de elementos complejos (trabajamos solo las reales por ahora) que tiene la característica de ser igual a su propia traspuesta conjugada (conjugado es, cambiarle el signo a la parte imaginaria). Es decir, el elemento en la  $i$ -ésima fila y  $j$ -ésima columna es igual al conjugado del elemento en la  $j$ -ésima fila e  $i$ -ésima columna, para todos los índices  $i$  y  $j$ .

En resumen, para matrices reales, que sea hermitiana es equivalente a que sea simétrica.

### 3.1.4 Matriz definida positiva

Una matriz es definida positiva si es hermitiana y todos sus autovalores son reales y positivos.

## 3.2 Método de Gauss

El método de Gauss para resolver el sistema  $Ax = b$  tiene dos partes; la primera es la triangularización del sistema, es decir, por medio de operaciones elementales, se construye un sistema  $A'x = b'$ , equivalente al primero, tal que  $A'$  sea triangular superior. La segunda parte es simplemente la solución del sistema triangular superior.

## 3.3 Método de Gauss con pivoteo parcial

En el método de Gauss clásico, únicamente se intercambian filas cuando el pivote,  $a_{k,k}$ , es nulo o casi nulo. Como el pivote (el elemento  $a_{k,k}$  en la iteración  $k$ ) será divisor para el cálculo de  $l_{ik}$ , y como el error de redondeo o de truncamiento se hace mayor cuando el divisor es cercano a cero, entonces es muy conveniente buscar que el pivote sea grande en valor absoluto. Es decir, hay que evitar los pivotes que sin ser nulos son cercanos a cero.

## 3.4 Método de Gauss-Jordan

Aquí, en vez de terminar en el momento que llegamos a una matriz triangular superior, seguimos con el método hasta llegar a la identidad. Queda así directamente resuelto el sistema de ecuaciones.

## 3.5 Factorización LU

La factorización LU es una forma de factorización de una matriz como el producto de una matriz triangular inferior (L) y una superior (U). Para realizar esta factorización la matriz debe ser no singular (invertible) pues esto nos garantiza su unicidad.

La idea es descomponer el sistema en dos subsistemas más sencillos de resolver:

Dada la ecuación matricial

$$Ax = LUx = b$$

Queremos la solución para un determinado A y b. Los pasos son los siguientes:

- Primero, resolvemos  $Ly = b$  para y
- Segundo, resolvemos  $Ux = y$  para x.

Nótese que ya tenemos las matrices L y U. La ventaja de este método es que es computacionalmente eficiente, porque podemos elegir el vector b que nos parezca y no tenemos que volver a hacer la eliminación de Gauss cada vez.

### 3.5.1 Doolittle y Crout

Llamamos descomposición de Doolittle al caso en que la matriz L es triangular inferior unitaria, es decir, que  $l_{i,i} = 1$ ,  $\forall i = 1, n$ .

Llamamos descomposición de Crout al caso en que la matriz U es triangular superior unitaria, es decir, que  $u_{i,i} = 1$ ,  $\forall i = 1, n$ .

## 3.6 Cholesky

Este método sirve para resolver el sistema  $Ax = b$  cuando la matriz A es definida positiva. A puede ser descompuesta como

$$A = LL^T$$

donde L es una matriz triangular inferior con entradas diagonales estrictamente positivas y  $L^T$  representa la conjugada traspuesta de L.

Se puede solucionar  $Ax = b$  calculando primero la descomposición de Cholesky  $A = LL^T$ , luego resolviendo  $Ly = b$  para y, y finalmente resolviendo  $L^Tx = y$  para x.

## 4 Sistemas de ecuaciones lineales - Métodos iterativos

### 4.1 Definiciones

#### 4.1.1 Normas vectoriales - matriciales

El concepto de norma corresponde simplemente a la abstracción del concepto de tamaño de un vector. Consideremos el vector que va de  $(0, 0, 0)$  a  $(2, 3, -4)$ . Su tamaño o magnitud es simplemente (euclidiana)  $\sqrt{2^2 + 3^2 + (-4)^2}$

Una norma es una función  $\mu : V \rightarrow R$  tal que:

- $\mu(x) \geq 0, \forall x \in V$
- $\mu(x) = 0 \Leftrightarrow x = 0$
- $\mu(\alpha x) = |\alpha|\mu(x), \forall \alpha \in R, \forall x \in V$
- $\mu(x + y) \leq \mu(x) + \mu(y), \forall x, y \in V$

Utilizamos las normas para medir qué tan lejos estamos de la solución buscada (se va comparando la norma de la solución en un momento dado con el error prefijado; cuando la norma es menor al error, se para).

Normas comunes:

- Norma 1 :  $\sum_{i=1}^n |v_i|$
- Norma 2 (euclidiana) :  $(\sum_{i=1}^n v_i^2)^{1/2}$
- Norma  $\infty$  :  $\max_{i=1,n} |v_i|$

Una norma es matricial si, además de las propiedades usuales de una norma, para cualquier par de matrices A y B,  $\|AB\| \leq \|A\| \cdot \|B\|$ .

Una norma matricial inducida por una norma vectorial se define de varias maneras, todas ellas equivalentes:

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

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

#### 4.1.2 Diagonal dominante

Una matriz es de diagonal estrictamente dominante, cuando lo es por filas o por columnas.

- Lo es por filas cuando, para todas las filas, el valor absoluto del elemento de la diagonal de esa fila es estrictamente mayor que la suma de los valores absolutos del resto de elementos de esa fila.
- Lo es por columnas cuando, para todas las columnas, el valor absoluto del elemento de la diagonal de esa columna es estrictamente mayor que la suma de los valores absolutos del resto de elementos de esa columna.

#### 4.1.3 Radio espectral

Sea A una matriz, el radio espectral  $\rho(A)$  es el máximo de los valores absolutos de los autovalores de A.

## 4.2 Método de Gauss-Seidel

Este método puede aplicarse a cualquier sistema de ecuaciones lineales que produzca una matriz cuadrada de coeficientes con los elementos de su diagonal no-nulos. En cada iteración del método de Gauss-Seidel, hay  $n$  subiteraciones. En la  $i$ -ésima subiteración se modifica únicamente  $x_i$ . Las demás coordenadas no se modifican. El cálculo de  $x_i$  se hace de tal manera que se satisfaga (en la  $k$ -ésima iteración):

$$x_i^k = \frac{b_i - (\sum_{j=1}^{i-1} a_{i,j} x_j^k + \sum_{j=i+1}^n a_{i,j} x_j^{k-1})}{a_{i,i}}$$

Matricialmente debemos tratar de escribir la matriz  $A$  como la suma de una matriz triangular inferior, una diagonal y una triangular superior  $A = (L + D + U)$ ,  $D = \text{diag}(a_{i,i})$ . Haciendo los despejes necesarios escribimos el mtodo de esta forma

$$x^{(k+1)} = -(L + D)^{-1} U x^{(k)} + (L + D)^{-1} b$$

por lo tanto  $M = -(L + D)^{-1} U$  y  $c = (L + D)^{-1} b$ . Decimos que  $M$  es la matriz de iteración del método y se puede escribir de la siguiente forma:

$$x^{(k+1)} = M x^{(k)} + c$$

La convergencia del mtodo solo se garantiza si:

- **Teorema 1:** Si  $A$  es de diagonal estrictamente dominante por filas, entonces el método de Gauss-Seidel converge para cualquier  $x_0$  inicial.
- **Teorema 2:** Si  $A$  es definida positiva, entonces el método de Gauss-Seidel converge para cualquier  $x_0$  inicial.

## 4.3 Método de Jacobi

Este método se parece al método Gauss-Seidel, también se utiliza la ecuación  $i$ -ésima para calcular  $x_i$  y el cálculo de  $x_i$  se hace de la misma forma. Pero un valor recién calculado de  $x_i$  no se utiliza inmediatamente. Los valores nuevos de  $x_i$  solamente se empiezan a utilizar cuando ya se calcularon todos los  $n$  valores  $x_i$ .

$$x_i^k = \frac{b_i - (\sum_{j=1, j \neq i}^n a_{i,j} x_j^{k-1})}{a_{i,i}}$$

Matricialmente la sucesión se construye descomponiendo la matriz del sistema  $A$  en la forma siguiente:

$$A = D + L + U$$

donde:

- $D$ , es una matriz diagonal.
- $L$ , es una matriz triangular inferior.
- $U$ , es una matriz triangular superior.

Partiendo de  $Ax = b$ , podemos reescribir dicha ecuación como:

$$Dx + (L + U)x = b$$

Luego,

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

Si  $a_{i,i} \neq 0$  para cada  $i$ . Por la regla iterativa, la definición del Método de Jacobi puede ser expresado de la forma:

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

Convergencia:

- Si la matriz  $A$  es estrictamente diagonal dominante. Además, puede converger incluso si esta condición no se satisface. (**condición suficiente**)
- Si el radio espectral ( $\rho$ ):  $\rho(D^{-1}R) < 1$ , siendo  $R = L + U$ . (**condición necesaria y suficiente**)

## 4.4 Método de Sobrerrelajación

Este método, conocido como SOR se puede considerar como una generalización del método Gauss-Seidel. En el método únicamente cambia la asignación, introduciendo un parámetro  $\omega$ :

$$x_i^k = \frac{\omega}{a_{i,i}}(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^k + \sum_{j=i+1}^n a_{i,j}x_j^{k-1}) + (1-\omega)x_i^{k-1}$$

Condiciones de convergencia:

- $0 < \omega < 2$  (**condición necesaria**).
- Si  $A$  es definida positiva, entonces converge para cualquier  $\omega \in (0, 2)$

## 5 Aproximación de autovalores

### 5.1 Definiciones

### 5.2 Teorema de Gerschgorin

El teorema de Gerschgorin es utilizado en álgebra lineal para encontrar una cota de los autovalores de una matriz compleja (o real) de orden  $n \times n$ .

Dada  $A$ , se definen los círculos  $D_1, \dots, D_n$  con centro en  $a_{ii}$  y radio  $r_i = \sum_{j=1, j \neq i}^n |a_{ij}|$ , el teorema afirma que los autovalores de la matriz  $A$  se encuentran en la unión de los  $n$  círculos. Además, cada componente conexa de esa unión contiene tantos autovalores como círculos haya en ella, donde círculos y autovalores son contados con multiplicidad.

### 5.3 Método de las potencias

Es un método iterativo que calcula sucesivas aproximaciones a los autovectores y autovalores de una matriz. Se usa principalmente para calcular el autovector de mayor autovalor (en valor absoluto) asociado.

Para aplicar el método de las potencias se supone que la matriz  $A$  de  $n \times n$  tiene  $n$  autovalores  $\lambda_1, \lambda_2, \dots, \lambda_n$  con un conjunto asociado de autovectores linealmente independientes  $(v^{(1)}, v^{(2)}, \dots, v^{(n)})$ . Es más, se supone que  $A$  tiene exactamente un autovalor  $\lambda_1$  cuya magnitud es la mayor, por lo que  $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \dots \geq |\lambda_n| \geq 0$ .

Requisito: al escribir el vector inicial en la base  $\lambda_1, \lambda_2, \dots, \lambda_n$ , el coeficiente de  $\lambda_1$  no debe ser nulo.

El método converge lentamente y sólo puede determinar uno de los autovectores de la matriz (el asociado al autovalor de mayor valor absoluto).

En cada paso  $k$ , se calcula  $x_{k+1} = \frac{Ax_k}{\|Ax_k\|}$ .

Este método puede usarse también para calcular el radio espectral de una matriz.

## 6 COSAS QUE NOS FALTAN

1. método de la falsa posición
2. factorización  $A = QR$
3. orden de convergencia