Universidad Simón Bolívar.
Informe de proyecto: proyecto 1.
Luis Diaz, 15-10420.
CO3211, Prof.: Saúl Buitriago.
Septiembre-Diciembre 2018

Demostración del teorema:

Sea el sistema [*]:

$$Az = y$$

Con A invertible y de solución única, queremos demostrar que se cumple la siguiente desigualdad:

$$\frac{\|z\|}{\|y\|} \leq \|A^{-1}\|$$

Como vimos en clase, la siguiente desigualdad [**] es verdadera, es una propiedad de las normas vectoriales :

$$\|Ax\| \leq \|A\| * \|x\|$$

Teniendo eso en cuenta y despejando de [*]:

$$Az = y$$

$$z = A^{-1} y$$

$$\textit{Por la propiedad } [**]: \quad \|z\| \leq \|A^{-1}\| * \|y\|$$

$$\frac{\|z\|}{\|y\|} \leq \|A^{-1}\|$$

□

Observaciones:

En el primer procedimiento se resolvía el sistema $Az = y$ para vectores y generados aleatoriamente. Esto se hizo con una multitud de vectores, y se observó que las aproximaciones más precisas eran aquellas asociadas a un vector cuyas componentes respetaban un cierto patrón de signos aparentemente aleatorio pero muy específico; los valores numéricos de dichas componentes no parecían tener tanta inceidencia como la tenían los patrones de signos. Bajo esta observación, se empieza a analizar el procedimiento 2, que realizaba la misma operación pero con vectores y de 1 y -1 posicionados aleatoriamente en las componentes de y.

Para el segundo procedimiento se hizo la observación de que los patrones de signos en las componentes del vector y tienen mucha incidencia y los resultados mejoraron enormemente, especialmente para la condición en norma infinito. La norma infinito mejoró bastante más su precisión frente a la norma 1; la primera se acercaba más a su equivalente nativa de matlab de lo que se acercaba la segunda.

En el tercer procedimientos se observaron resultados casi exactos en todas las matrices salvo en la matriz A4 y en A5. Se presume que se debe a la gran cantidad de 0's en sus componentes y el tamaño de la matriz. En el resto de casos la aproximación fue completamente exacta con la condición en norma 1 de matlab/octave.

Conclusiones:

Teniendo en cuenta las observaciones realizadas en cada una de las pruebas, decidimos buscar cuál era el patrón de signos en el vector aleatorio que generaba mejores resultados, pues los vectores que ofrecían mejores resultados siempre compartían un patrón de signos muy específico para ser aleatorio. Lo primero fue notar lo siguiente:

Sea el sistema:

$$Az=y$$

$$z=A^{-1}y$$

Aquí podemos observar que la componente i de z se obtiene como el producto de vectores entre la fila i de A^{-1} y el vector y. Y más aún, si el vector y fuera de solo 1, entonces ese producto sería realmente la suma de los elementos de esa fila de A^{-1}. Si esa suma fuera con el valor absoluto de los elementos de la fila, y además la mayor de las sumas de las filas, entonces en z_i estaría el valor de la norma infinito de A. Así, si además de ser un vector de 1, el vector y tuviera signos negativos en la posición y_i si A_{ij} es negativo, y signos positivos de lo contrario, entonces todos los elementos en la suma de esa fila serían positivos. Bajo esta premisa, intentamos averiguar en qué posiciones es necesario poner un negativo para guardar la suma en valor absoluto.

Supongamos que el mejor de los resultados en el procedimiento 2 está asociado a una combinación apropiada de signos; en ese caso en una de las componentes de z estaría directamente la norma inf de la matriz A , esto podría explicar porqué se obtuvieron resultados tan buenos para este caso. Sin embargo, si la norma inf está en la compontente z_i que a su vez se obtuvo de la suma en valor absoluto de la fila, esto quiere decir que en la fila i de A^{-1} es la que tiene la mejor suma con el valor absoluto de sus elementos, es decir, la fila utilizada para calcular la norma infinito. Con esto en mente es que pasamos a analizar el procedimiento 3.

El procedimiento 3 es más complejo, para empezar resuelve el sistema $Ax=b$, por nuestro análisis anterior podemos apreciar que x_i contiene la suma de la fila ${A^{-1}}_i$. Luego, el procedimiento decide construir un vector b tal que $b_i=-1$ si $x_i<0$ y $b_i=1$ de lo contrario. ¿Por qué escoger los signos de esta forma?. Se cree que esto se debe a que como x_i almacena la suma de la fila ${A^{-1}}_i$ multiplicada por la constante positiva $1/n$, si esa suma es negativa, es *más probable* que el elemento ${A^{-1}}_{ij}$ con mayor valor absoluto sea negativo, y si la suma es positiva, entonces *es más probable* que ese elemento sea positivo. Al hacer esto, tenemos una combinación aproximada bastante buena a la combinación de signos que tiene la columna de A^{-1} cuya norma 1 es la mayor.

Hecho esto, se resuelve el siguiente sistema:

$$A^t z=b$$

Y aquí observamos lo siguiente:

$$z=(A^t)^{-1} b$$

$$z=(A^{-1})^t b$$

Ahora, z_i ya no tiene la suma de los valores absolutos de la fila i de A^{-1} , sino de $(A^{-1})^t$. Pero, ¿Quién es esta fila? Es claro que al ser la traspuesta de la inversa, esta fila realmente es una columna de A^{-1} . En este sentido, y por como se configuró el vector b , en algún z_i estará guardada la suma de la columna de A con la mayor norma 1. Digamos que está en z_i , esto quiere decir que la columna i de A^{-1} es aquella que tiene la mayor norma 1, es por eso que luego se construye otro sistema nuevo de la siguiente forma:

$$Ax=y$$

Donde $y_i=1$ si z_i es el máximo en valor absoluto de z , y $y_i=0$ de lo contrario. Este nuevo vector y es la columna de la identidad asociada a la columna de A^{-1} con mayor norma 1, y como vimos en clase, se puede obtener la columna j de la inversa resolviendo un sistema $Ax=b$ si b es la columna

j de la identidad. Como y es la columna i de la identidad, el sistema planteado obtiene la columna i de la inversa que sabemos que tiene la mayor norma 1 de las columnas de la inversa. Así, es claro que $\|x\|_1 \approx \|A^{-1}\|_1$ (es la mejor aproximación a $\|A^{-1}\|_1$. Como ya tenemos $\|A\|_1$, el condicionamiento lo calculamos como $\|x\|_1 \|A\|_1$.

En conclusión, este procedimiento busca primero el índice j que tiene más probabilidades de identificar a la columa de la matriz inversa cuya norma 1 sea la mayor de todas las columnas, y una vez obtenido ese índice, busca el vector en cuestión y calcula su norma.

Tabla de resultados:

Matrices	**Método1**		**Método2**		**Método3**	**Condición real**	
	n-1	**n-inf**	**n-1**	**n-inf**	**n-1**	**n-1**	**n-inf**
A1	2833.2	3253.9	2260.5	4488.0	4488.0	4488.0	4488.0
A2	282.70	282.02	202.41	510.94	1277.5	1277.5	950.38
A3	460.99	661.34	493.89	2039.8	4657.2	4657.2	5747.5
A4	NaN	NaN	NaN	NaN	NaN	992795.05	572038.71
A5	18.775	50.686	17.131	86.291	76.062	114.69	98.336

Vemos que para las matrices 4 y 5 los métodos no funcionan tan bien. En el caso de la matriz 4 se puede deber tanto a la dimensión de la matriz como a la naturaleza de los valores de esta; todos los números son 0 o valores cercanos a 0, así que pueden ocurrir divisiones entre números muy pequeños lo que llevaría a resultados demasiado grandes para ser representados.