Facultad de Matemática y Computación (UH)

Ciencia de la Computación

Matemática Numérica

Curso 2021

Conferencia 4

1. Lo que deberías saber de la actividad anterior

El profesor llega al aula, saluda a los presentes y comienza a hablar.

—En la última actividad tuvieron muchas, muchas, muuuuuuchas actividades, pero que les dieron la oportunidad de deducir varios resultados importantes. Aquí va un resumen de lo mínimo que deberían saber después de eso.

1.1. Matrices eficientes

—La primera actividad era identificar tipos de matrices que te permitieran resolver el sistema $Ax = b$ en menos de $O(n^2)$ operaciones. A ese tipo de matrices les vamos a llamar matrices eficientes.

`Matrices eficientes:`

—Dentro de estas matrices eficientes debían identificar, al menos, las diagonales, las triangulares inferiores y superiores, y las ortogonales.

```
Matrices eficientes:

  ▪ diagonales,

  ▪ triangulares inferiores,

  ▪ triangulares superiores,

  ▪ ortogonales.
```

—Si la matriz del sistema es diagonal, se resuelve en $O(n)$ dividiendo por los elementos de la diagonal. Si es triangular, se pueden usar dos algoritmos que se llaman de sustitución hacia adelante para las inferiores, y de sustitución hacia atrás para las superiores. Finalmente las matrices ortogonales son aquellas que su inversa es su traspuesta...

$$\texttt{M es ortogonal si } \mathtt{M}^{-1} = \mathtt{M}^{t}$$

—Y si A es ortogonal, el sistema $Ax = b$ se puede resolver como $x = A^t b$. Por otro lado, también pudieron deducir uno las 10 mejores ideas del siglo XX en materia de cálculo científico.

Escribe en la pizarra

```
Descomposición PLU
```

1.2. Descomposición PLU

—Toda matriz A, cuadrada y no singular, se puede descomponer en el producto de tres matrices P, L y U...

```
A = PLU
```

Rocio:

En algunos casos, la descomposición es $PA = LU$, y que en dependencia de cuál de los dos se use, la matriz P se construye de manera diferente. Eso es bueno tenerlo claro cuando se va a leer esto en algún libro, porque si no, se te arman tremendos líos, rollos y confusiones.

—En el caso de la descomposición PLU, P es una matriz de permutaciones, L es una matriz triangular inferior con 1 en la diagonal, y U es una matriz triangular superior, que es justamente la matriz que se obtiene al aplicar el método de Gauss.

```
P:  Matriz de permutación
L:  Triangular inferior con 1 en la diagonal
U:  Triangular superior, resultado de aplicar Gauss.
```

—Además, algunas personas demostraron que las matrices L y U son únicas, y que el elemento que está en la posición (i, j) de L es el multiplicador usado para lograr un 0 en la posición (i, j) de U.

Se vira, y ve algun@s alumn@s asintiendo con cara de "¡Yo sí lo hice!", mientras otros tienen cara de "Ah, mira de donde salen esos números". Cuando tiene una idea de cuántos hicieron los ejercicios del taller, sigue hablando.

—La importancia de la matriz P tiene que ver con las estrategias de pivote.

1.3. Estrategias de pivote

—Las estrategias de pivotes permiten evitar errores numéricos en el cálculo de L y U. Algunas de las más usadas son el pivote parcial, pivote parcial escalado y pivote total.

- `Pivote parcial`
- `Pivote parcial escalado`
- `Pitove total`

Deja de hablar y mira a los estudiantes. No ve ninguna cara de pánico ni de ¿y eso qué cosa es? así que sigue hablando.

—El pivote parcial consiste en seleccionar como pivote al mayor elemento de la columna que se está analizando. El pivote parcial escalado consiste en escoger el mayor elemento de la columna, pero tomando en cuenta cuán grande es con respecto a los demás elementos de su fila. Por último la estrategia de pivote total consiste en escoger como pivote el mayor elemento de toda la submatriz que queda por analizar.

- `Pivote parcial: máximo de la columna.`
- `Pivote parcial escalado: máximo, en comparación con su fila`
- `Pitove total: máximo de todos los elementos restantes.`

> **Arnel:**
> En el caso de usar la estrategia de pivote total, la descomposición no es $A = PLU$, sino $A = PLUQ$, donde Q es otra matriz de permutación, pero como está multiplicando a la derecha, lo que hace es intercambiar columnas.

El profesor mira a los estudiantes, y ve que hay varios asintiendo con cara de "Sí, eso también lo descubrí yo".

—Si eso lo tienen claro, podemos pasar a la forma numéricamente correcta de resolver un sistema de ecuaciones $Ax = b$.

1.4. Solución de un SEL usando PLU

—Una vez que se tiene la factorización PLU de una matriz A, la mejor forma de resolver el sistema de ecuaciones lineales $Ax = b$ es sustituir A por su descomposición:

$$\mathtt{Ax = b} \iff \mathtt{PLUx = b}$$

—Y lo primero es resolver el sistema...

$$\mathtt{Ly = P^{-1}b}$$

> **Rocio:**
> Ahí no se está cometiendo ninguna maldición imperdonable porque las matrices de permutación son matrices ortogonales. Además, por el efecto que tiene multiplicar por una matriz de permutación, ni siquiera hay que realizar la multiplicación, porque a partir de vector de permutaciones que se obtiene al aplicar el método de Gauss para no hacer los intercambios de fila, se pueden realizar los cambios correspondientes al vector b.

—Una vez que se conoce el valor de `y`, después de aplicar la sustitución hacia adelante a este sistema —*señala al sistema* $Ly = P^{-1}b$—, se puede obtener el valor de x, al aplicar la sustitución hacia atrás a:

$$\mathtt{Ux = y}$$

—Esos dos sistemas de ecuaciones lineales se resuelven en $O(n^2)$.

$$\texttt{Ly = P}^{-1}\texttt{b: Sustitución hacia adelante.}$$

$$\texttt{Ux = y: Sustitución hacia atrás.}$$

—¿Alguna duda hasta aquí?

Como la mayoría de las personas había hecho estos ejercicios, no hay dudas.

—Una de las aplicaciones más frecuentes que tiene la descomposición PLU es que permite resolver eficientemente varios sistemas de ecuaciones lineales que todos tengan la misma matriz.

1.5. Solución de varios SELs con la misma matriz A

—Cuando se quiere resolver varios sistemas de ecuaciones lineales con la misma matriz A...

$$\begin{aligned}\texttt{Ax} &= \texttt{b}_1\\ \texttt{Ax} &= \texttt{b}_2\\ &\vdots\\ \texttt{Ax} &= \texttt{b}_\texttt{n}\end{aligned}$$

—lo que se hace es calcular la factorización PLU de A una sola vez, y después aplicar sustitución hacia adelante y hacia atrás con las matrices P, L y U cambiando los términos independientes.

```
A=PLU (Una sola vez)
Para cada i:
  Calcular y, solución de: Ly = P^{-1}b_i
  Obtener x_i, solución de Ux_i = y
```

Rocio:

Aunque casi nunca hace falta, en caso de necesidad extrema, se puede calcular la inversa de una matriz, resolviendo n sistemas de ecuaciones lineales con la misma matriz A. Los términos independientes de esos sistemas son los vectores canónicos del espacio $\mathbb{R}^n$.

—¿Alguna pregunta hasta aquí?

Silencio.

1.6. Final del resumen

—Si no tienen ninguna pregunta hasta aquí, pueden revisar su kit de supervivencia de la asignatura.

Cuando el profesor termina de hablar aparece el cartel Kit de Superveviencia Numérica en el aire delante de tus ojos, a unos 20 centímetros de distancia. Miras a las demás personas, pero no ves ningún otro cartel. Sin embargo, sí ves que casi todo el mundo está levantando una mano y hace el gesto de apretar un botón imaginario que parece estar en el aire, a unos 20 centímetros de sus ojos. Por la reacción que ellos, cuando presionaron el botón pasó algo. Decides presionar tu botón, y comienzas a levantar la mano derecha.

Carlos:

Si quieren, pueden apretar el botón con la mano pero no hace falta. Es suficiente imaginarse que lo aprietan para que funcione. Hacerlo con la mano es un poco lento, pero tiene la ventaja de todos se ven muy graciosos haciendo ese gesto. Con los botones apropiados en las posiciones correctas puede salir una coreografía interesante... pero no hace falta. Solo imagínense que lo presionan y ya.

Bajas la mano derecha y presionas el botón mentalmente. Del texto Kit de Superveviencia Numérica se despliegan tres elementos: Serie de Taylor, Épsilon de la Máquina y Descomposición PLU.

Carlos:

Si quieren, pueden abrir el menú de problemas. No se asusten. Son los problemas que ya pueden resolver usando numérica. Háganlo. En ese menú hasta solo tienen un botón: SELs. Si lo presionan —*varios alumnos lo presionan*—, ven las herramientas que tienen para resolver Sistemas de Ecuaciones Lineales. Si presionan ese botón (no hace falta que lo hagan con la mano) todo el mundo puede ver la descomposición PLU. Sin embargo, en dependencia de lo que hayan hecho en el taller, algunas personas pueden ver también la factorización de Cholesky, que permite descomponer una matriz A simétrica y definida positiva en el producto de dos matrices L y L^t. Si no has hecho esa actividad del taller, deberías, porque a veces, para este tipo de matrices, la descomposición PLU puede ser demasiado lenta.

—Con esos elementos ya estamos listos para...

Se vira para escribir en la pizarra, pero una muchacha de la primera fila lo interrumpe.

—Profe, ¿cómo se quita el cartelito? Es que tenerlo delante todo el tiempo es incómodo.

El profesor se da vuelta y le dice:

—De la forma tradicional. Presiona la crucecita de cerrar. En Ajustes puedes configurarlo todo para que te sea cómodo.

Arnel:

Antes de que pregunten, por defecto, Ajustes está en la parte derecha superior. No tienen que tocarlo con la mano. Con imaginarlo, ya aparece.

—Gracias, profe.

Cierras el Kit de supervivencia, y solo para verificar vuelves a "imaginar" Kit de supervivencia. El cartel aparece nuevamente flotando a unos 20 centímetros de tus ojos. ¡Funciona! Lo cierras y le prestas atención al profesor, que está a punto de comenzar a hablar, pero una voz de varón lo interrumpe.

—Profe, a mí no me aparecen los Ajustes.

El profesor lo mira extrañado.

—Lo único que tienes que hacer es imaginarte que salen los ajustes y ya aparecen.

—No profe, me lo imagino y no aparece.

El profesor mira al resto del aula.

—¿Todo el mundo pudo ver sus ajustes?

Todos los alumnos, menos el muchacho, dicen que sí. El profesor lo vuelve a mirar fijamente, con el ceño fruncido, y de pronto le dice.

—Vamos a hacer una prueba rápida... Imagínate Ajustes... sin h.

—Ah, ya. ¡Gracias profe! ¡Ya salió!

Toda el aula mira al alumno, que está moviendo las manos delante de él poniendo la configuración a su gusto. El profesor se vira hacia a los profesores de clase práctica y les dice.

—Creo que deberíamos incorporar otro libro a la bibliografía.

Arnel dice que él se encarga y comienza a escribir algo. El profesor se vira hacia los alumnos y comienza a hablar.

—Ya que la descomposición PLU está incorporada al kit de supervivencia numérica, y la tienen como una opción para resolver sistemas de ecuaciones lineales, estamos listos para conocer a uno de los mayores problemas e inconvenientes de toda el Álgebra Lineal Numérica. Causa tantos problema que hay quienes lo comparan con un supervillano. ¿Están listos para eso?

La mayoría de los alumnos dicen que sí, que están listos para saber cuál es uno de los mayores inconvenientes del Álgebra Lineal Numérica. El profesor sonríe y empieza a hablar.

—Tener esa actitud es muy bueno, pero en realidad solo están listos para empezar a estar listos —*sonríe*—. Necesitamos un par de elementos más. El primero, se supone que sí lo deben saber y solo se trata de refrescarlo es la norma de un vector.

2. Normas de vectores

El profesor escribe en la pizarra:

```
Norma de un vector x: ||x||
```

—Una norma es una función de Rn en R que cumple cuatro propiedades: siempre es mayor que cero, es cero solo si su argumento es 0, la norma de un escalar por un vector es el módulo del escalar por el vector, y finalmente, cumple la desigualdad triangular.

Cuando termina de hablar se puede ver lo que escribió:

Norma $||\cdot|| : \mathbb{R}^n \to \mathbb{R}$, tal que:

1. $||\mathtt{x}|| \geq 0$

2. $||\mathtt{x}|| = 0 \iff \mathtt{x} = 0$

3. $||\alpha\mathtt{x}|| = |\alpha|||\mathtt{x}||$

4. $||\mathtt{x}+\mathtt{y}|| \leq ||\mathtt{x}|| + ||\mathtt{y}||$

—Algunas de las normas más conocidas son la norma 2, la norma 1 y la norma infinito. La norma 2 es la raíz cuadrada de la suma de los cuadrados de los elementos del vector. La norma 1 es la la suma de los módulos y la norma infinito es el máximo de los módulos de los elementos.

Termina de escribir al mismo tiempo que termina de hablar, y en la pizarra está escrito lo siguiente:

Si $\mathtt{x} \in \mathbb{R}^{\mathtt{n}}$:

$$||\mathtt{x}||_2 = \sqrt{\mathtt{x}_1^2 + \mathtt{x}_2^2 + \ldots \mathtt{x}_{\mathtt{n}}^2}$$

$$||\mathtt{x}||_1 = |\mathtt{x}_1| + |\mathtt{x}_2| + \ldots |\mathtt{x}_{\mathtt{n}}|$$

$$||\mathtt{x}||_\infty = \max_{\mathtt{i}=1..\mathtt{n}} |\mathtt{x}_{\mathtt{i}}|$$

Mira a los alumnos por si hay que explicar algo más, pero todos están muy tranquilitos tomando notas. Es como si se lo supieran, o como si no se lo supieran pero no tuvieran ningún interés en que el profesor se entere de ese detalle.

—Por ejemplo, si x es el vector [1,2]

```
x = [1, -2]
```

—la norma 2 es $\sqrt{5}$.

$$||\mathtt{x}||_2 = \sqrt{1^2 + (-2)^2} = \sqrt{5}$$

—La norma 1 es 3,

$$||\mathtt{x}||_1 = |1| + |-2| = 3$$

—y la norma infinito es 2.

$$||\mathtt{x}||_\infty = \text{máx}\{|1|, |-2|\} = 2$$

—¿Alguna pregunta hasta aquí?

No hay preguntas.

—Si no hay preguntas, entonces podemos reescribir las definiciones de errores usando normas, para poder aplicarlas también cuando estemos trabajando con vectores.

Carlos:

En realidad no es reescribirlas, las definiciones son casi las mismas. Incluso las que tienen módulo son las mismas, porque el módulo es una norma en $\mathbb{R}$.

—Si x es un vector con el valor real, y $\overline{x}$ es una aproximación de x, se pueden definir el error absoluto y relativo —*comienza a escribir en la pizarra*—...

Sheila:

Siempre recuerden que la definición de error absoluto no varía: es el verdadero valor menos el aproximado. Los ejemplos que hemos puesto hasta ahora han sido con números, pero esa definición se aplica a lo que sea. Si tienes un elefante real y un elefante aproximado, el error absoluto se define como el elefante real menos el elefante aproximado.

Y no importa lo que les hayan dicho sus profesores de cuarto grado con respecto a eso de sumar o restar elefantes. Con un poquito de imaginación y las operaciones adecuadas se puede formar un espacio vectorial donde los elementos sean elefantes y ya ahí puedes restarlos sin problemas.

Cuando el profesor termina de escribir se puede leer:

$$\begin{aligned} \mathtt{e} &= ||\mathtt{x} - \overline{\mathtt{x}}|| \\ \delta &= \frac{||\mathtt{x} - \overline{\mathtt{x}}||}{||\mathtt{x}||} \end{aligned}$$

Rocio:

En este caso, la única forma que hay de calcular el error relativo es usando las normas así que, por suerte, en esta conferencia no hay necesidad de poner el ejemplo de los pozuelos.

—Por ejemplo, si el verdadero vector es $x = [1, 2]$, y se quiere aproximar por el vector $\overline{x} = [0.8, 2.3]$, el error absoluto sería:

$$\mathtt{e = || x - \overline{x}|| = ||[0.2,-0.3]|| = 0.3}$$

—Ahí se usó la norma infinita, porque es la que menos cálculos tiene.

Gustavo:

En todos los sistemas de cálculo científico también hay una función que te permite calcular la norma de un vector. En casi todos los casos, te permite especificar qué norma quieres usar, y si no se especifica nada, usa una por defecto. En la mayoría de estos sistemas, la función se llama `norm`.

—En ese mismo caso, el error relativo sería:

$$\delta = \frac{e}{||x||} = \frac{0.3}{2} = 0.15$$

—en este caso —*señala al error relativo*—, también está calculado con la norma infinito.

Arnel:

Es importante, cuando se vaya a calcular el error absoluto usar la misma norma en el numerador y el denominador.

Rocio:

Cuando se trabaja con vectores, los errores absolutos y relativos pueden variar en dependencia de la norma, pero como todas las normas son equivalentes, la selección de la norma para calcular los errores no suele ser demasiado importante... salvo algunos casos, en los que sí lo es. Pero cuando eso pase, lo vamos a dejar muy claro. Así que se puede asumir que para calcular los errores absolutos y relativos de cálculos con vectores, se puede usar cualquier norma.

2.1. Errores en la solución de SELs

—Cuando se trabaja la solución de Sistemas de Ecuaciones Lineales, además del error absoluto y relativo, hay otra magnitud que suele ser relevante. Se llama residual y se define como —*comienza a escribir en la pizarra*—.

```
Ax = b:  Sistema de Ecuaciones Lineales
     x̄:  aproximación
     r:  residual que se define como:
                r = b - Ax̄
```

—El residual —*señala la expresión del residual que está en la pizarra*— se puede interpretar como "cuánto le falta a $A\overline{x}$ para ser igual al vector b".

Rocio:

El residual es un vector. Si queremos saber "de qué tamaño" es el residual, calculamos la norma de r y ¡listo!.

Y ya que saben lo que es el residual y cómo calcular la norma de un vector, es hora de hablar de un ejercicio de la clase práctica 0: el que te preguntan cómo se puede saber si un vector dado es solución de un sistema de ecuaciones lineales. Muchas personas usaron el operador `==` para comparar los vectores, y otras usaron la función `np.array_equal`, que también usa el operador `==`.

Arnel:

No se preocupen, en aquella época ustedes eran jóvenes e inexpertos, y no sabían que no lo podían usar, pero más allá de los problemas numéricos que eso tiene, hay otro problema más: y es que como se comparan los elementos de los vectores término a término solo se puede determinar si los vectores son iguales o no. Hay una alternativa que permite saber no solo si son iguales, sino que además permite saber, en caso de que sean diferentes, cuán diferentes son.

Esa alternativa es calcular la norma del residual $||r|| = ||b - A\overline{x}||$. Por la segunda propiedad de la norma el residual tendría norma cero solo si $A\overline{x}$ es exactamente igual a b, y en caso de que sea diferente de cero, el valor de esa norma te da una idea de cuán diferentes son.

Carlos:

Como estamos trabajando en una aritmética de punto flotante, y se realizan varias operaciones casi nunca ese residual será exactamente 0, así que la forma numéricamente correcta de determinar si un vector es solución de un sistema de ecuaciones lineales es comprobar si la norma del residual es lo suficientemente pequeña. ¿Qué significa lo suficientemente pequeña? Depende del problema, pero si no hay ningún requerimiento especial, se puede asumir un valor entre 10^{-10} y 10^{-12}.

—¿Alguna duda hasta aquí? Es que se ha hecho tremendo silencio de pronto y es como si no hubiera nadie haciéndome caso.

Los alumnos dejan de leer lo que dicen los profesores de clases prácticas y se concentran en el profesor de conferencia.

—Además del residual, también se puede definir el residual relativo —*comienza a escribir en la pizarra*—.

Sheila:

El residual se puede ver como una interpretación del error absoluto, y por lo tanto, se calcula igual: el verdadero valor menos el aproximado. ¿Cuál es el verdadero valor en este caso? b. ¿Cuál es el valor aproximado? $A\overline{x}$.

Y si el residual es un error absoluto, que se calcula igual y todo, ¿cómo se calcularía el residual relativo? Exactamente igual que el error relativo: el absoluto sobre el verdadero valor. Y en el caso del residual, ya sabes cuál es el verdadero valor, ¿verdad?

Sheila termina de "hablar" al mismo tiempo que el profesor termina de escribir. Cuando en la pizarra se puede ver:

$$\texttt{residual relativo: } \frac{||\texttt{r}||}{||\texttt{b}||}$$

—La ventaja que tiene el residual relativo es que, al igual que el error relativo, debe ser independiente del tamaño de los vectores. Con esto terminamos la primera parte de los requisitos necesarios para llegar al supervillano. ¿Alguna duda hasta aquí?

Nadie tiene dudas.

—Perfecto.

Al mismo tiempo que el profesor habla, a los alumnos que estaban prestando atención les aparece una notificación:

ADQUIRIDAS LAS SIGUIENTES HERRAMIENTAS:

- NORMA DE VECTORES.

- Error absoluto para vectores.
- Error relativo para vectores.
- Residual.
- Residual Relativo.

Herramientas pendientes para tener acceso al supervillano: 1.

3. Normas de matrices

—Si no hay dudas con las normas de los vectores, podemos introducir el último elemento que necesitamos: las normas de matrices, que se parecen mucho a las de vectores. Una norma de matriz es cualquier función de $\mathbb{R}^{n\times n} \to \mathbb{R}$ que cumple las mismas cuatro propiedades de la norma de vectores:

$$||\cdot|| : \mathbb{R}^{n\times n} \to \mathbb{R} \texttt{ tal que:}$$

1. $||\mathtt{A}|| \geq 0$
2. $||\mathtt{A}|| = 0 \iff \mathtt{A} = 0$
3. $||\alpha\mathtt{A}|| = |\alpha|||\mathtt{A}||$
4. $||\mathtt{A}+\mathtt{B}|| \leq ||\mathtt{A}|| + ||\mathtt{B}||$

—y además satisface una propiedad extra: la norma del producto es menor o igual que el producto de las normas.

$$||\mathtt{AB}|| \leq ||\mathtt{A}||||\mathtt{B}||$$

—Por lo tanto, una norma de matrices es cualquier función $\mathbb{R}^{n\times n} \to \mathbb{R}$ que satisfaga todas estas propiedades —*se para al lado de la pizarra y señala a lo que está escrito*—:

$||\cdot|| : \mathbb{R}^{n\times n} \to \mathbb{R}$ `tal que:`

1. $||\mathtt{A}|| \geq 0$
2. $||\mathtt{A}|| = 0 \iff \mathtt{A} = 0$
3. $||\alpha\mathtt{A}|| = |\alpha|||\mathtt{A}||$
4. $||\mathtt{A}+\mathtt{B}|| \leq ||\mathtt{A}|| + ||\mathtt{B}||$
5. $||\mathtt{AB}|| \leq ||\mathtt{A}||||\mathtt{B}||$

Se queda en silencio esperando que los alumnos terminen de copiar.

Sheila:

Al igual que las normas de vectores, todas las normas de matrices son equivalentes, así que se puede usar cualquiera de ellas indistintamente, a no ser que alguna propiedad o teorema necesite alguna norma específica, que en ese caso lo vamos a dejar bien claro. Pero mientras no digamos nada, pueden usar la que más cómoda les sea.

—Las normas de matrices más frecuentes son la norma infinito, la norma 1 y la norma 2.

$$||\cdot||_\infty,\ ||\cdot||_1,\ ||\cdot||_2$$

—La norma infinito de una matriz es el máximo de la suma de los módulos de los elementos de cada fila —*hace una pausa*—, y eso que dije, aunque es correcto, no se entiende nada —*hace otra pausa*—. Pero no se preocupen, cuando se escribe sí se entiende.

Comienza a escribir en la pizarra.

Sheila:

Es verdad que cuando se escribe es mejor. De hecho, cualquier cosa es mejor que eso que él dijo, pero igual no se hagan muchas ilusiones con entenderlo rápido.

Cuando termina de escribir señala a la pizarra donde se puede ver:

$$\texttt{Si A} \in \mathbb{R}^{n\times m}\texttt{:}$$
$$||\texttt{A}||_\infty = \max_{\texttt{i=1..n}} \left\{ \sum_{\texttt{j=1}}^{\texttt{m}} |\texttt{a}_{\texttt{ij}}| \right\}$$

—Ahora sí. La norma infinito de una matriz es el máximo por fila de la suma de los módulos de los elementos de la fila.

Sheila:

Como les había dicho, no se hagan ilusiones. La buena noticia es que ahora pone un ejemplo y ahí sí todo queda claro.

—Vamos a poner un ejemplo —*comienza a escribir*—.

Sheila:

Lo que él ha tratado de decir las dos veces es que hay que ir fila por fila sumando los módulos de los elementos, para después quedarse con la mayor de esas sumas. Con el ejemplo debe quedar claro.

—Vamos a verlo con esta matriz —*señala a lo que acaba de escribir*—.

$$\mathtt{A} = \begin{pmatrix} 1 & -3 & -2 \\ -4 & 5 & 3 \\ 8 & 3 & 7 \end{pmatrix}$$

—Para la norma infinito, necesitamos la suma de los módulos de los elementos de cada fila. Vamos a llamar f_i a la suma de los módulos de la fila i.

$$\begin{array}{ccccccccc} \mathtt{f}_1 & = & |1| & + & |-3| & + & |-2| & = & 6 \\ \mathtt{f}_2 & = & |-4| & + & |5| & + & |3| & = & 12 \\ \mathtt{f}_2 & = & |8| & + & |3| & + & |7| & = & 18 \end{array}$$

—Y ahora, de esos valores, nos quedamos con el mayor que es 18, y esa es la norma infinito de la matriz A.

$$||\mathtt{A}||_\infty = 18 = \mathtt{max}\,\{\mathtt{f}_1, \mathtt{f}_2, \mathtt{f}_3\} = \mathtt{max}\{6, 12, 18\}$$

—¿Alguna duda hasta aquí?

> **Sheila:**
>
> Si quieres ahora, con el ejemplo resuelto, puedes verificar que la definición original es correcta y que tiene sentido. O que no es correcta y hay algo que arreglar, y en ese caso puedes ganar créditos extras.

Como nadie responde, sigue hablando.

—La norma 1 es similar a la norma infinito, pero tomando el máximo por columnas en vez de por filas: es el máximo de la suma de los módulos por columna, que se puede escribir matemáticamente como:

$$||\mathtt{A}||_1 = \max_{\mathtt{j=1..n}} \left\{ \sum_{\mathtt{i=1}}^{\mathtt{n}} |\mathtt{a_{ij}}| \right\}.$$

Sheila:

Igual que con la norma infinito, aquí se entiende mejor con el ejemplo, pero en esencia, es lo mismo que la norma infinito, pero ahora por columnas.

—Para calcular la norma 1 de la misma matriz A —*señala donde está escrito*—:

$$\mathtt{A} = \begin{pmatrix} 1 & -3 & -2 \\ -4 & 5 & 3 \\ 8 & 3 & 7 \end{pmatrix}$$

—necesitamos la suma de los módulos de los elementos en la columna i, que vamos a denotar por c_i.

$$\begin{array}{lllllllll} \mathtt{c}_1 & = & |\ 1| & + & |-4| & + & |-8| & = & 13 \\ \mathtt{c}_2 & = & |-3| & + & |\ 5| & + & |\ 3| & = & 11 \\ \mathtt{c}_2 & = & |-2| & + & |\ 3| & + & |\ 7| & = & 12 \end{array}$$

—Y por tanto, la norma 1 de la matriz A es 13.

$$||\mathtt{A}||_1 = 13 = \mathtt{max}\,\{\mathtt{c}_1, \mathtt{c}_2, \mathtt{c}_3\} = \mathtt{max}\{13, 11, 12\}$$

—¿Alguna duda hasta aquí?

No hay dudas.

—Entonces solo nos queda la norma 2, que es un poquito más costosa de calcular, porque se define como la raíz del mayor valor propio de A traspuesta por A:

$$||\mathtt{A}||_2 = \sqrt{\lambda_{max}(\mathtt{A}^t\mathtt{A})}$$

—donde λ_{max} representa el mayor valor propio.

Gustavo:

En todos los sistemas de cálculo científico como Numpy en Python, MATLAB u Octave, hay una función que calcula los valores propios de una matriz. Esa función casi siempre se llama `eig`, o algo similar, porque en inglés los valores propios se llaman *eigenvalues*.

Arnel:

De hecho, si están trabajando con algún sistema de cálculo científico que no tenga un comando para calcular los valores propios, deberían cambiarlo, o por lo menos, no deberían decir que están trabajando con un sistema de cálculo científico.

—En el caso de la matriz A:

$$\mathtt{A} = \begin{pmatrix} 1 & -3 & -2 \\ -4 & 5 & 3 \\ 8 & 3 & 7 \end{pmatrix}$$

—la norma 2 es aproximadamente 11.17.

$$||\mathtt{A}||_2 \approx 1:124.7689$$

—y los cálculos no los vamos a hacer aquí.

Mira a los alumnos por si alguien quiere protestar porque no se vayan a hacer los cálculos, pero todos parecen muy felices con la idea.

Carlos:

De todas formas, no estaría mal que verificaran que ese es el valor correcto de la norma. Los valores propios deben estar bien calculados porque se calcularon con Octave, pero nada garantiza que el profesor no haya cambiado algún elemento de la matriz o algo así... Con él nunca se sabe, y es preferible comprobar las cosas. Además, quizás te puedas ganar algunos créditos por eso.

Gustavo:

En todos los sistemas de cálculo científico también hay una función que te permite calcular la norma de una matriz o un vector. En casi todos los casos, te permite especificar qué norma quieres usar, y si no se especifica nada, usa una por defecto. En la mayoría de estos sistemas, la función se llama `norm`.

—¿Alguna duda? ¿Alguien quiere preguntar algo?

Silencio.

—Entonces, con las normas de matrices ya estamos listos para presentar al primer supervillano de la matemática numérica —*hace un silencio*—... aunque por el nombre, creo que es una supervillana. Independientemente del género que tenga, vamos a coger los 5 minutos.

Mientras te pones de pie para salir del aula puedes ver un nuevo mensaje:

ADQUIRIDA LA HERRAMIENTA:

- NORMA DE MATRICES.

ACCESO AL SUPERVILLANO DESBLOQUEADO.

Cierras el mensaje y sales para disfrutar de los 5 minutos.

4. Residual y error

—Con las normas de matrices ya estamos listos para presentar uno de los mayores problemas de toda el álgebra lineal numérica, y es el siguiente. Tenemos un sistema de ecuaciones lineales $Ax = b$, y una aproximación de la solución $\overline{x}$.

```
Ax=b  Sistema de Ecuaciones Lineales
x̄     Solución aproximada
```

—¿Cómo podemos saber cuán buena es $\overline{x}$ como aproximación de la solución real?

> **Rocio:**
>
> Antes de que alguien proponga usar el error absoluto o relativo, recuerden que para calcular cualquiera de esos dos necesitamos el verdadero valor, y aquí no lo tenemos. Por eso, cualquier solución que propongan no debe incluir a esos errores.

Un par de alumnos que habían levantado la mano muy rápido, después de leer lo que dijo Rocio, la bajan más rápido todavía. Al mismo tiempo, una muchacha sentada en la segunda fila habla sin levantar la mano.

—Profe, ¿y ahí no se pudiera usar el residual, que se acaba de introducir en esta conferencia?

—¡Muy bien! Anótate 10 000 créditos por una muy buena idea —*la muchacha comienza anotarse los créditos*—... que no funciona.

La muchacha deja de escribir y mira al profesor con cara de ¿cómo que eso no funciona?

—Pero igual anótate los créditos, que la idea es buena, y funcionaría muy bien, si no fuera por el problema que vamos a ver ahora.

La alumna termina de anotarse los créditos, mientras el resto del aula espera que el profesor diga cuál es el problema.

—Si les parece, les propongo analizar primero por qué es una buena idea, y después, por qué no funciona. ¿Están de acuerdo?

> **Arnel:**
> Esa última pregunta es retórica. Él lo va hacer en ese orden independientemente de lo que digamos, así, respondan lo que quieran, o no respondan nada... da lo mismo.

Algunos alumnos dicen que sí, otros se quedan callados, y otros se encongen de hombros mientras dicen Ahhhhloeehehagana.

—Perfecto. Entonces, vamos a hacerlo en ese orden. Primero, por qué es una buena idea. Eso es fácil. Si $\overline{x}$ es una buena aproximación de la solución, entonces Ax debe ser un vector "parecido" a b.

$$\texttt{Si } \overline{\texttt{x}} \approx \texttt{x, entonces}$$
$$\texttt{A}\overline{\texttt{x}} \approx \texttt{Ax = b.}$$

—Y si $A\overline{x}$ es un vector "parecido" a b, entonces la norma del residual $r = b - A\overline{x}$ debe ser pequeña.

$$\texttt{Si } \overline{\texttt{x}} \approx \texttt{x, entonces A}\overline{\texttt{x}} \approx \texttt{Ax = b}$$

`y por tanto ||r|| = ||b - A`$\overline{\texttt{x}}$`|| debe ser pequeña.`

—Y por otro lado, mientras menor sea ese residual, mejor debería ser $\overline{x}$ como aproximación de la solución del sistema, porque —*comienza a escribir en la pizarra*—:

`Si ||r|| = ||b - A`$\overline{\texttt{x}}$`|| es pequeño,`

`entonces ||Ax - A`$\overline{\texttt{x}}$`|| es pequeño`

`y entonces ||x - `$\overline{\texttt{x}}$`|| debería ser pequeño.`

—Y todos deberíamos estar de acuerdo con eso, ¿no?

Algunos alumnos asienten sin mucha seguridad, porque ya el profesor dijo que eso no funciona... pero la verdad es que si eso no funciona, voy a perder el poquito de fe que tenían en la humanidad `:-(`.

—Perfecto, si todos estamos de acuerdo...

Un alumno desde la segunda fila.

—Profe, permiso. A mí me preocupa una cosa. Por todo eso que usted está diciendo, ¿también nos vamos a tener que pelear con los profesores de Álgebra? Digo, porque si eso no está bien —*señala lo que está escrito en la pizarra*—, creo que ellos se van a molestar un poco, y con razón.

El profesor la mira.

—Creo que, afortunadamente, eso no va a ser un problema, sobre todo porque ellos saben lo peligroso que puede ser asumir que las propiedades matemáticas se cumplen cuando estamos trabajando con las computadoras.

Rocio:

Una de las ventajas de pasar el curso de Matemática Numérica es que se interioriza que la Matemática es una cosa, y lo que pasa en las aritméticas de punto flotante es otra. Por eso, siempre se debe tener mucho cuidado con lo que se programa, y con lo que se asume.

—En este caso, la propiedad que estamos asumiendo que se cumple es que al resolver un sistema de ecuaciones lineales —*comienza a escribir en la pizarra*—, si el residual es pequeño, entonces el error es pequeño, y al revés.

$$\texttt{Residual pequeño} \iff \texttt{Error pequeño}$$

—Esto tiene mucho sentido, sobre todo en el caso extremo en que el error sea cero. Ahí sí no hay discusión: en ese caso, el residual tiene que ser cero también, porque si tenemos la solución exacta, Ax tiene que ser igual a b.

$$\mathtt{e = ||x - \overline{x}|| = 0} \iff \mathtt{\overline{x} = x}$$

$$\texttt{y por lo tanto: } \mathtt{A\overline{x} = Ax = b.}$$

—¿Todos estamos de acuerdo con eso?

Todo el mundo asiente, incluso l@s escéptic@s, porque ahí no hay nada con lo que estar en desacuerdo.

—El problema es que hay veces que eso no pasa —*todos los alumnos abren los ojos con cara de ¿Qué?*—, y además, ya ustedes lo experimentaron una vez.

El profesor mira a los alumnos y sonríe mentalmente, porque todos los alumnos se ven igualitos que el español de Elpidio Valdés que dice ¿Yooooo?

—Sí, ustedes. Clase Práctica 0, ejercicio 9. En ese ejercicio ustedes multiplican la matriz A por la solución "exacta" del sistema, y el resultado que obtienen no es b, Y mientras más grande sea la matriz, menos se parece Ax a b y por lo tanto mayor es la norma del residual.

Hace una pausa para que los estudiantes asimilen lo que se acaba de decir.

> **Sheila:**
>
> Quizás aquí sea bueno que aproveches esta pausa para comprobar que en efecto, en ese ejercicio, cuando n mayor que 60 se tiene un error de exactamente 0, y un residual mayor que 0.

Algun@s alumn@s, l@s más list@s, dejan de leer esto y van a comprobar los resultados de ese ejercicio de la CP 0.

Cuando los alumnos terminan de comprobar que, en efecto, con un error 0 se puede tener un residual mayor que 0 miran a la pizarra y ven que está escrito:

`El error puede ser pequeño (¡incluso 0!) y el residual grande.`

Cuando todos los alumnos terminan de hacer la comprobación (el profesor se da cuenta de que hicieron la comprobación por la cara que tienen) empieza a hablar.

—Veo que ya lo comprobaron, pero esa es solo en una dirección. También puede pasar al revés. El residual puede ser pequeño y el error grande. Eso lo pueden comprobar en el ejercicio 2 de la CP 4 que es la de esta semana.

El profesor deja de hablar.

Sheila:

Aquí el profesor va a esperar hasta que hagas el ejercicio. Realmente vale la pena hacerlo, sobre todo porque es un muy buen ejemplo de un caso donde el residual es pequeño (del orden de 10^{-4}) y el error es 10000 veces mayor que eso. Pero a mí no me hagan caso, que a lo mejor esos números están mal. Mejor compruébenlo ustedes.

Cuando todos los estudiantes terminan de hacer el ejercicio 2 de la CP de esta semana, el profesor sigue hablando.

—La conclusión de todo esto es que el residual y el error pueden no tener nada que ver el uno con el otro.

Rocio:

Si alguien se está preguntando cuál es el problema con eso, solo tiene que recordar que el residual es la única forma que tenemos nosotros para comprobar cuán bueno es $\overline{x}$ como aproximación de la verdadera solución. En otras palabras, lo que quiere decir la conclusión es que la única forma que tenemos para verificar si nuestra solución es buena, no sirve para verificarlo.

—Esa es la mala noticia. Por suerte tenemos dos buenas noticias. La primera es que eso no siempre ocurre, y la segunda es que hay formas de saber cuándo sí se cumple. Para saber cuándo vamos a tener ese problema necesitamos dejar varias cosas por escrito —*comienza a escribir*—.

Se quiere resolver: $Ax = b$ (1)

donde, solución real: x

solución aproximada: $\overline{x}$

error absoluto: $e = x - \overline{x}$ (2)

residual: $r = b - A\overline{x}$ (3)

Termina de escribir, señala a la pizarra y pregunta:

—¿Alguna duda con eso?

Nadie tiene dudas.

—Perfecto, entonces vamos a hacer algunas manipulaciones algebraicas, que cuando las terminemos vamos a tener la explicación de por qué a veces el residual y el error pueden no tener nada que ver.

Comienza a escribir en la pizarra mientras habla.

—Lo primero es darnos cuenta de que el residual es en realidad la matriz multiplicada por el error. Para eso, solo hay que darse cuenta de que b es Ax, y se tiene... esto.

Se vira y señala lo que acaba de escribir.

$b = Ax$, por lo que

$r = b - A\overline{x}$ se puede escribir como

$r = Ax - A\overline{x}$. Si se saca factor común A

$r = A(x - \overline{x})$, y por definición de e,

$r = Ae$

—¿Alguna duda con eso?

Silencio.

—Si no hay dudas, podemos incorporarlo a los resultados que tenemos.

```
 Se quiere resolver:  Ax = b      (1)
donde, solución real:  x
  solución aproximada:  x̄
      error absoluto:  e = x - x̄   (2)
            residual:  r = b - Ax̄  (3)
            residual:  r = Ae      (4)
```

—El siguiente paso puede ser peligroso, y es —*se vira hacia la pizarra y escribe mientras habla*— expresar e en términos de la inversa de A...

Algunos alumnos abren mucho los ojos cuando oyen la palabra inversa. Otros sonríen mientras piensan: "Perfecto, en cuanto calcule la inversa... ¡lo suspendo! Ellos dijeron que calcular la inversa es una maldición imperdonable y el que lo haga por gusto, suspende automáticamente... ¡Siempre he querido suspender a un profesor!"

Rocio:

En este caso, el uso de la inversa no es un problema, porque no la vamos a calcular numéricamente. Solo lo vamos a hacer en el papel. Este es uno de esos casos en los que se puede usar la inversa sin problema.

La decepción se refleja en las caras de algunos estudiantes. El profesor termina de escribir, se vira y señala lo que está escrito.

Se tiene que: r = Ae   (4)

Y de ahí: $e = A^{-1}r$

—Este resultado también lo podemos incorporar a los que ya tenemos:

Se quiere resolver:	$Ax = b$	(1)
donde, solución real:	x	
solución aproximada:	$\bar{x}$	
error absoluto:	$e = x - \bar{x}$	(2)
residual:	$r = b - A\bar{x}$	(3)
residual:	$r = Ae$	(4)
error absoluto:	$e = A^{-1}r$	(5)

—¿Alguna pregunta hasta aquí?

Claro que no hay preguntas, si lo único que ha hecho ha sido escribir cosas que ya sabíamos o que eran muy fáciles de deducir a partir de lo que ya sabíamos.

—Perfecto. ¿Qué tenemos hasta ahora? Se puede tener la solución exacta de un sistema de ecuaciones lineales y el residual puede ser mayor que 0. Por otro lado, se puede tener un residual muy pequeño y el error puede ser grande. También tenemos que el residual es la matriz A por el error *—señala donde está escrito eso en la pizarra—*, y que el error se puede obtener como A^{-1} multiplicada por el residual.

Rocio:

También tenemos que ese cálculo de la inversa no es una maldición imperdonable, porque no la estamos calculando numéricamente.

—Por último, tenemos las propiedades de las normas, en particular que la norma del producto es menor o igual que el producto de las normas. Con eso, ya podemos presentar al mayor villano de toda el álgebra lineal numérica. Vamos a empezar.

Continuará...

5. Condición de una matriz

—Lo que vamos a hacer es mostrar que el error relativo está acotado por el residual relativo, multiplicado por otro número. Y ese número es el que nos interesa, es el que causa los mayores problemas dentro del álgebra lineal numérica, y que incluso ha inspirado boleros.

Rocio:
Para aquellas personas a las que les resulte extraña esta idea de un tema musical inspirado por un concepto de análisis numérico, puede escuchar el bolero Usted, de José Antonio Zorrilla Martínez, cuya versión más famosa es la que interpreta Luis Miguel en su disco Romance de 1991. Esa canción completa, está dedicada a este concepto. Las dos primeras estrofas están dedicadas literalmente a este número y todos lo problemas que causa. Para que no tengan problemas con la música, se lo dejamos en el EVEA.

—Lo que sigue es un juego algebraico donde se aplica la propiedad multiplicativa de las normas, y un poco de ingenio, a partir de esto —*señala lo que está escrito en la pizarra*—:

$$
\begin{aligned}
\texttt{Se quiere resolver:}\quad & \texttt{Ax = b} && (1)\\
\texttt{donde, solución real:}\quad & \texttt{x} && \\
\texttt{solución aproximada:}\quad & \overline{\texttt{x}} && \\
\texttt{error absoluto:}\quad & \texttt{e = x - }\overline{\texttt{x}} && (2)\\
\texttt{residual:}\quad & \texttt{r = b - A}\overline{\texttt{x}} && (3)\\
\texttt{residual:}\quad & \texttt{r = Ae} && (4)\\
\texttt{error absoluto:}\quad & \texttt{e = A}^{-1}\texttt{r} && (5)
\end{aligned}
$$

—Si a partir de ahora hay algún paso que no les quede claro, me avisan —*comienza a escribir mientras sigue hablando*—. Lo primero que hay que hacer es aplicar la norma a ambos lados de (5):

$$||e|| = ||A^{-1}r||$$

—Y por la propiedad multiplicativa de las normas tenemos eso es menor o igual que:

$$||e|| \leq ||A^{-1}|| \; ||r||$$

—Ahora, el término derecho de esa desigualdad lo voy a multiplicar y dividir por la norma de b, para obtener el residual relativo:

$$||e|| \leq ||A^{-1}|| \; ||r|| \; \frac{||b||}{||b||}$$

—Y eso se puede escribir como:

$$||e|| \leq ||A^{-1}|| \; ||b|| \; \frac{||r||}{||b||}$$

—Y fíjense que tenemos una cota para el error que depende de la norma de la inversa de A, de b, y del residual relativo —*mientras menciona cada uno de esos elementos los señala en la pizarra*—. ¿Alguna duda hasta ahí? *No hay dudas.*

—En realidad, la cota del error no depende de la norma de b, porque por esto —*señala en la pizarra donde está escrito $Ax = b$*—, se tiene que b es Ax, por lo que tendríamos

$$\texttt{||e||} \leq \texttt{||A}^{-1}\texttt{||} \ \texttt{||Ax||} \ \frac{||\texttt{r}||}{||\texttt{b}||}$$

—pero $||Ax||$ es menor o igual que el producto $||A||\ ||x||$:

$$\texttt{||Ax||} \leq \texttt{||A||} \ \texttt{||x||}$$

—Por lo tanto, el error absoluto es menor o igual que:

$$\texttt{||e||} \leq \texttt{||A}^{-1}\texttt{||} \ \texttt{||A||} \ \texttt{||x||} \ \frac{||\texttt{r}||}{||\texttt{b}||}$$

—Y como estamos hablando del residual relativo —*señala la fracción de norma de x sobre norma de b*—, es más apropiado compararlo con el error relativo y no con el error absoluto, y eso es justamente lo que obtenemos si dividimos toda la expresión por la norma de x.

$$\frac{||\texttt{e}||}{||\texttt{x}||} \leq \texttt{||A}^{-1}\texttt{||} \ \texttt{||A||}\frac{||\texttt{r}||}{||\texttt{b}||}$$

—¿Y qué tenemos hasta aquí? La explicación de por qué se puede tener un residual relativo pequeño y un error grande. Fíjense que si el residual relativo es pequeño —*señala al residual relativo*—, pero este producto es grande —*señala la norma de la inversa de A por la norma de A*—, el error puede ser grande.

Sheila:

Fíjense que el error relativo puede ser tan pequeño como ustedes quieran, por ejemplo menor que el épsilon de la máquina (10^{-16}), que si el producto de la norma de A por la norma de la inversa de A es mayor que el recíproco del épsilon de la máquina, por ejemplo 10^{17} ya el error relativo puede ser de tamaño 10, y un error relativo de tamaño 1, ya es un problema. Imagínense uno de tamaño 10.

Gustavo:

En todos los sistemas de cálculo científico hay una función que te permite calcular la condición de una matriz, y casi siempre esa función tiene un nombre parecido a `cond`. Para tener una idea de cuál puede ser el problema con el ejercicio 9 de la clase práctica 0, cuando la matriz tiene una dimensión mayor que 60, ayuda calcular el producto de la norma de A, por la norma de la inversa de A. De hecho, ahora voy a entretener al profesor un momento para que puedas hacer ese cálculo.

Gustavo se acerca al profesor y le dice algo muy bajito. El profesor le responde también muy bajito, y Gustavo asiente, pero cuando el profesor se calla, él vuelve a decirle algo, que nadie oye, nada más que ellos dos. Ahora quien asiente es el profesor... ¿Y tú qué haces leyendo esto? `:-/` *¡Se supone que Gustavo está entreteniendo al profesor para que tú puedas calcular el producto de la norma de A por la norma de su inversa! Deja "el chisme" y saca las cuentas esas.*

Rocio:

Aquí sí vas a cometer una maldición imperdonable, porque vas a calcular la inversa de una matriz... Pero no te preocupes. Hay veces que no no queda más remedio que calcularla. Este no es uno de esos casos, pero no importa. Para este ejemplo puntual, puedes hacerlo sin preocupación. Además, el profesor no se va a enterar porque está entretenido hablando con Gustavo. Y si se entera, tampoco importa porque este es un caso sirve de ejemplo para ilustrar los problemas que puede tener ese número.

Sheila:

Fíjense que en ese caso, a pesar de que el error es 0, el residual comienza a ser diferente de 0, cuando $||A||||A^{-1}||$ es mayor que el recíproco de el épsilon del a máquina.

Mientras habla con el profesor, Gustavo está constantemente mirando a los alumnos, hasta que varios de ellos le hacen un gesto indicando que ya, que ya hicieron los cálculos. Gustavo le dice una última cosa al profesor y se sienta. El profesor mira a los alumnos, mira a la pizarra, y empieza a hablar.

—Lo fundamental de este resultado...

Señala donde está escrito:

$$\frac{||\mathtt{e}||}{||\mathtt{x}||} \leq \mathtt{||A^{-1}||}\ \ \mathtt{||A||}\frac{||\mathtt{r}||}{||\mathtt{b}||}$$

—es que hay veces en las que no hay garantía de que un residual relativo pequeño sea indicador de un error pequeño. ¿De qué depende que lo sea o no? Del valor de $||A^{-1}||\ ||A||$. Ese valor tiene un nombre y es **Condición de la matriz**.

$$\texttt{Condición de una matriz A: cond(A) = ||A|| ||A}^{-1}\texttt{||}$$

—Y por lo tanto, se puede reescribir la desigualdad como:

$$\frac{||\mathtt{e}||}{||\mathtt{x}||} \leq \mathtt{cond(A)}\frac{||\mathtt{r}||}{||\mathtt{b}||}$$

—Fíjense que si la condición de la matriz es grande, el residual relativo puede ser un muy mal indicador del error relativo. Pero, si la condición de la matriz es un valor cercano 1, entonces el residual relativo puede ser una cota superior para el error relativo. Por lo tanto, siempre que sea posible, conviene trabajar con matrices que estén bien condicionadas. Eso significa que su condición esté cercana a 1.

Carlos:

No pregunten qué pasa si la condición de una matriz es menor que 1. Eso nunca puede pasar. La explicación es un ejercicio sencillito de la clase práctica.

—Lo malo de este resultado...

Señala nuevamente la desigualdad:

$$\frac{||\mathtt{e}||}{||\mathtt{x}||} \leq \mathtt{cond(A)}\frac{||\mathtt{r}||}{||\mathtt{b}||}$$

—es que, literalmente, es solo la mitad del problema, porque en realidad, la desigualdad tiene una segunda parte...

Comienza a escribir en la pizarra. Cuando termina, se quita y se puede ver:

$$\frac{1}{\mathtt{cond(A)}}\frac{||\mathtt{r}||}{||\mathtt{b}||} \leq \frac{||\mathtt{e}||}{||\mathtt{x}||} \leq \mathtt{cond(A)}\frac{||\mathtt{r}||}{||\mathtt{b}||}$$

—Y ahora sí tenemos el problema completo. Fíjense que si condición de la matriz es 1, entonces el error relativo y el residual relativo son exactamente iguales y el mundo es una maravilla, y las matrices y las soluciones de sus sistemas pueden vivir felices por siempre jamás. Pero, en la medida que la condición de una matriz comienza a crecer, la diferencia entre el residual relativo y el error relativo puede ser cada vez mayor, y por lo tanto, no se puede confiar en el residual para tener una idea del error. Se pueden tener residuales pequeños y errores grandes —*señala la primera desigualdad*—, o errores pequeños y residuales grandes —*señala la segunda desigualdad*—.

Sheila:

En realidad, aunque él está hablando de errores y residuales, lo que aparecen en las desigualdades son errores y residuales relativos, pero casi siempre esos problemas se trasladan a los absolutos también, así que es mejor pensar que en cualquiera de los dos casos (absolutos o relativos) puede haber problemas.

Arnel:

En otras palabras, y quitando algunos tecnicismos, la mayor implicación de que una matriz A esté mal condicionada es que no puedes confiar en los resultados de resolver un sistema de ecuaciones donde la matriz del sistema sea A, porque nunca podrás saber si están correctos o no. Eso es importante recordarlo siempre.

Rocio:

Y ese es el motivo porque el cual la condición de una matriz es un gran problema, en toda el álgebra lineal numérica. Cuando aparece una mala condición, no se puede confiar en los resultados.

—¿Alguna pregunta hasta aquí?

Al parecer, no hay preguntas.

—Entonces, si no hay preguntas vamos a ver uno de los inconvenientes de la condición de una matriz, y es que en la definición de la condición aparece la inversa de A —*señala a la inversa en la definición de la condición*—.

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

—Fíjense que calcular la inversa para obtener la condición no tiene mucho sentido, porque —*mira a los alumnos*—... ¿cómo se calcularía la inversa de una matriz?

Se queda esperando que alguien diga la respuesta, y varias personas dicen que resolviendo n sistemas de ecuaciones lineales.

—¡Exacto, y ese es el problema! Para calcular la inversa necesitas resolver sistemas de ecuaciones lineales, y ese resultado lo vas a usar para calcular el número de condición. Pero que la matriz esté mal condicionada significa precisamente que no se puede confiar en las soluciones de los sistemas de ecuaciones lineales con esa matriz. Así que si la matriz está mal condicionada, probablemente la condición que obtengas también esté mal. Por lo tanto, si la matriz está mal condicionada no lo puedes saber calculando la condición.

Carlos:

Si están pensando qué hacer si no puedes calcular la inversa, te digo que existen alternativas: una permite aproximar el número de condición de una matriz A resolviendo dos sistemas de ecuaciones lineales con la misma matriz A, y el otro permite calcular la condición de la matriz usando la norma 2, a través de la factorización SVD de la matriz A. En ninguno de los dos casos es necesario calcular la inversa.

—La conclusión de todo esto es que, al igual que hay que tener en cuenta que los cálculos que hagas con una aritmética de punto flotante pueden tener errores que dependen únicamente de la aritmética, cuando resuelvas sistemas de ecuaciones lineales hay que considerar que pueden aparecer errores que dependen únicamente de la condición de la matriz, y que puede ser la causa de muchos problemas y dolores de cabeza.

Comienza a sonar bajito la canción Usted, interpretada por Luis Miguel.

Rocio:

Como curiosidad histórica, entre las primeras personas que se dieron cuenta de que la condición de una matriz podía ser un problema estuvieron Alan Turing y John Von Neumann, dos de los padres de las computadoras modernas. Eso es para que tengan una idea de que la condición de las matrices han estado ~~jodiendo~~ *causando problemas* incluso desde antes del surgimiento de las computadoras.

—Y con esta parte de la condición de una matriz terminamos el kit de supervivencia elemental para la solución de sistemas de ecuaciones lineales...

Aparece un nuevo menú Kit de problemas. *Cuando le das desplegar puedes ver:*

Kit de problemas:

- Matriz mal condicionada.

—Y ya con eso estamos listos para pasar a otros temas de la asignatura. ¿Alguna pregunta?

Nadie pregunta.

—Entonces, nos vemos la semana próxima.

Los alumnos recogen sus cosas y la canción "Usted" sube de volumen, mientras los alumnos salen del aula.

Bibliografía recomendada

- *Numerical Analysis.* Burden R. L., Faires J. D. y Annette M. Burden, 10th Edition. Brooks Cole Publishing, 2016.
- *Numerical Analysis.* Timothy Sauer, 3rd Edition. Pearson, 2017.
- *Accuracy and stability of numerical algorithms.* 2nd Edition. Nicholas J. Higham SIAM. 2002.
- *Elementary Numerical Analysis, An algorithmic approach.* S. D. Conte y Carl de Boor. 3rd Edition. McGraw-Hill Book Company. 1980.
- *Ortografía y Gramática para Dummies.* Pilar Comín Sebastián. Colección Para Dummies. 2018. ISBN: 978-84-329-0478-3.