



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico 1

“No creo que a él le gustará eso”

Metodos numericos
Primer Cuatrimestre de 2015

Integrante	LU	Correo electrónico
Gastón Zanitti	058/10	gzanitti@gmail.com
Ricardo Colombo	156/08	ricardogcolombo@gmail.com



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (54 11) 4576-3359

<http://www.fcen.uba.ar>

Contents

1	Introduccion teorica	3
2	Desarrollo	4
2.1	Matriz banda	4
2.2	Almacenamiento de la matriz banda	4
2.3	Eliminacion gaussiana sobre matriz banda	5
2.4	Descomposicion LU	5
2.5	No hay tiempo que perder	6
2.6	Implementacion Simple	6
2.7	Implementacion Sherman-Morrison	7
3	Discusion	8
3.1	Manipulación de estructura interna de la matriz	8
4	Resultados	9
4.1	Análisis de temperatura y tiempos en función de las discretizaciones	9
4.2	Optimización del cómputo para cambios leves del sistema	13
5	Conclusiones	14
5.1	Ventajas en el uso de una matriz banda	14
5.2	Discretización de un problema continuo	14
5.3	Pérdida de precisión en el uso de aritmética de punto flotante	14
5.4	Otras aplicaciones	14
6	Apendice	15

1 Introduccion teorica

En este trabajo practico intentaremos modelizar y resolver el problema de una superficie a la que se le aplica calor en ciertos puntos, teniendo como condiciones ademas que los bordes permanecen a temperatura constante. Para modelizar este problema utilizaremos la ecuacion del calor:

$$\frac{\partial^2 T(x, y)}{\partial x^2} + \frac{\partial^2 T(x, y)}{\partial y^2} = 0. \quad (1)$$

Con esta ecuacion diferencial, es posible calcular la temperatura en cualquier punto de la superficie. Sin embargo, dado que queremos resolver el sistema bajo un modelo que no sea continuo, será necesario discretizar la ecuacion diferencial de alguna manera adecuada. Para ello, discretizaremos esta superficie en segmentos de superficie, que afectaran la precision de las respuestas. Parece intuitivo pensar que mientras mas pequeños sean los segmentos, el sistema discreto mas se asemejará con el continuo, obteniendo así respuestas mas parecidas a este.

Aprovechando la discretizacion del sistema y que este es un problema lineal, lo modelizaremos como un problema $Ax = b$ sobre la cual aplicaremos diversas tecnicas, como eliminacion gaussiana o descomposicion LU, que nos permitiran resolver las incognitas de una manera mas comoda.

Para decirlo mas formalmente, dados a y b el ancho y el alto de nuestra superficie, respectivamente, h la granularidad con la que discretizaremos, y valiendo que $a = m \times h$ y $b = n \times h$, obtendremos una grilla de $(n+1) \times (m+1)$ puntos (donde el punto $(0, 0)$ se corresponde con el extremo inferior izquierdo).

Llamemos $t_{ij} = T(x_j, y_i)$ al valor (desconocido) de la función T en el punto $(x_j, y_i) = (jh, ih)$. La aproximación finita (que es posible gracias a la discretización realizada sobre el sistema) afirma que

$$t_{ij} = \frac{t_{i-1,j} + t_{i+1,j} + t_{i,j-1} + t_{i,j+1}}{4}. \quad (2)$$

De esta forma, es posible plantear un sistema en donde cada punto esté en función de otros y así resolver todas las ecuaciones nos dará la temperatura en el punto crítico.

2 Desarrollo

2.1 Matriz banda

Una vez obtenida la discretización de nuestro sistema y la posición de los puntos de calor en esta versión como se mencionó con anterioridad, planteamos las ecuaciones que nos permitirán resolverlo. Consideremos r al radio de alcance de los puntos de calor producidos por las sanguijuelas, T_c su temperatura y C_1, \dots, C_k los k puntos de calor, con $C_i \in R^2$, $C_i = (x_i, y_i)$.

Sabemos que si $x = 0 \vee y = 0 \vee x = m \vee y = n \Rightarrow T(x, y) = -100$.
También que si $\exists C_j, C_j = (x_j, y_j) / \sqrt{(x - x_j)^2 + (y - y_j)^2} \leq r \Rightarrow T(x, y) = T_c$.
Luego, nos falta definir $T(x, y)$ para todos los (x, y) que no son alcanzados por estas condiciones. Ahora plantearemos $n \times m$ ecuaciones con $n \times m$ incógnitas, que serán el sistema que luego resolveremos. Las incógnitas serán los $T(x, y)$, y sus constantes correspondientes a_{ij} .

Sea (α, β) un punto de la grilla:

1. Si es parte del borde, entonces su ecuación será $a_{ij} = -100$ por la temperatura por defecto que toma esta sección.
2. Si es un punto donde existe una sanguijuela, este toma el valor de la constante de temperatura otorgado por esta. $a_{ij} = k$ donde K representa la temperatura pasada como input para dicha sanguijuela.
3. Si en cambio es un punto del interior del plano que no se corresponde con la posición de ninguna sanguijuela, su ecuación será, como mencionamos con anterioridad,

$$t_{ij} = \frac{t_{i-1,j} + t_{i+1,j} + t_{i,j-1} + t_{i,j+1}}{4}. \quad (3)$$

Teniendo en cuenta esta información, podremos despejar la incógnita para introducir dicha ecuación dentro de nuestra matriz, de la siguiente manera:

$$0 = -4t_{ij} + t_{i-1,j} + t_{i+1,j} + t_{i,j-1} + t_{i,j+1}. \quad (4)$$

logrando así igualar la ecuación a cero.

Nótese que si cada fila de nuestra matriz de resolución representa las ecuaciones del sistema y los t_{ii} representan el valor puntual que representa dicha ecuación, entonces cada t_{ij} que forme parte de esta estará a una distancia x constante. Llevando esta idea al gráfico de la matriz, puede empezarse a entrever como, al completar estas ecuaciones, la matriz banda empieza a tomar forma.

Una vez obtenida esta matriz, ya estamos listos para comenzar con la resolución del sistema.

2.2 Almacenamiento de la matriz banda

Debido a que la representación matricial del problema que atacamos posee una estructura donde los elementos no nulos se concentran en la diagonal, pudimos representar esta matriz de $(n \times m) \times (n \times m)$ usando otra estructura de $n \times (2m + 1)$ elementos, sabiendo que más allá de la banda diagonal de $p \times q$ elementos, el resto de la matriz se completa con 0.

Inicialmente utilizamos un arreglo unidimensional para almacenar la estructura, amparandonos en una característica presente en gcc y c99 que permite definir el tamaño de un arreglo en tiempo de

ejecución. (Cuando lo corrimos en clang esto fallaba). Si bien esta estructura de datos es muy eficiente, decidimos movernos a la representación en un vector unidimensional, obteniendo mayor comodidad de manipulación. La diferencia en tamaño de la estructura del vector frente al arreglo es despreciable frente al almacenamiento general que requiere el problema.

2.3 Eliminacion gaussiana sobre matriz banda

Como primer forma de resolver el sistema de ecuaciones, se implemento un Gauss sobre el que se realizaron algunas mejoras. La idea general del algoritmo se mantiene, en el paso k diagonalizamos la columna k de la matriz A y todos los elementos en la columna k -ésima a partir del $a_{k,k}$ serán 0.

La mejora que implementaremos será que, sabiendo que la matriz A es banda, a partir del elemento $a_{k+p,k}$ (donde p es el límite de la banda) los siguientes serán 0 y a partir del elemento $a_{k,k+p}$ también serán 0.

Luego para la columna k no será necesario chequear hasta el final de la matriz, sino hasta el valor $k + p$.

Lo mismo pasa con las filas. En el gauss estándar, luego de diagonalizar una columna, todos los elementos de esa fila deben ser actualizados. Pero al trabajar con una matriz banda, solo es necesario actualizar los p valores que siguen a la diagonal, ya que tenemos asegurado que todos los valores siguientes serán 0.

Escrito de manera más formal el algoritmo será el siguiente:

TP1 1 void Gauss(matriz A, vector b)

- 1: Para $i=1\dots n$
 - 2: Tomo el elemento $a_{i,i}$ como pivote
 - 3: Para $k = i + 1, \dots i + p$
 - 4: Para $w = i + 1, \dots k + p$
 - 5: $a_{k,w} = a_{k,w} - a_{i,w} * (a_{k,i} / a_{w,w})$
 - 6: $a_{i,i} = 0$
-

De este algoritmo se desprende que este algoritmo tiene una mejor complejidad comparado con el Gauss estándar, siendo la misma de $O(n * p^2)$ siendo n la cantidad de incógnitas de nuestra matriz, y p el tamaño de la banda.

Una vez concluido el proceso de eliminación gaussiana sobre la matriz banda, utilizamos backwards substitution para resolver las ecuaciones y conseguir así el valor de cada uno de los puntos de calor del plano discretizado. Aquí también es posible realizar una optimización, ya que en cada paso no es necesario ir hasta la última columna sino hasta la columna donde sabemos que comienza la banda.

De esta manera habremos disminuimos la complejidad del backwards substitution de $O(n^2)$ a una menor de $O(n * p)$.

2.4 Descomposicion LU

La segunda forma que desarrollamos para resolver el problema es realizando una descomposición LU a la matriz original. La ventaja de esto es que en caso de actualizar el vector b , el costo de volver a obtener los resultados se reduce en $O(n^2)$.

Dado que la descomposición LU es muy similar al algoritmo de eliminación gaussiana, este también puede ser optimizado para ser utilizado en una matriz banda. Las optimizaciones serán similares a las de Gauss. Expresándolo de manera algorítmica, la descomposición LU optimizada será así:

TP1 2 void Gauss(matriz A, vector b)

```
1: Inicializo una matriz L con unos en la diagonal
2: Para i=1...n
3:   Tomo el elemento  $a_{i,i}$  como pivote
4:   Para  $k = i + 1, \dots, i + p$ 
5:     Para  $w = i + 1, \dots, k + p$ 
6:        $a_{k,w} = a_{k,w} - a_{i,w} * (a_{k,i}/a_{w,w})$ 
7:        $a_{i,i} = 0$ 
8:        $L_{i,i} = a_{i,w} * (a_{k,i}/a_{w,w})$ 
```

Luego en la matriz A obtengo la matriz diagonal superior y en el L la diagonal inferior.

Puede verse que siendo el algoritmo muy similar al de eliminacion gaussiana, tambien posee la misma complejidad, esto es $O(n * p^2)$.

Este algoritmo sera utilizado mas adelante, cuando queramos invertir una matriz de manera barata para poder utilizar el algoritmo de Sherman Morrison.

2.5 No hay tiempo que perder

En esta parte del trabajo, buscaremos la posibilidad de disminuir la temperatura del punto critico a travez de la eliminacion de una sanguijuela. Dado que esto nos afecta el sistema de ecuaciones, en primera instancia, será necesario recalcular todo el sistema otra vez, lo que podria resultar muy costoso. Plantear otra vez las ecuaciones, ahora con esta variación, y volver a calcular las incógnitas mediante el algoritmo de eliminación gaussiana tomaría otra vez complejidad $O(n * p^2)$. Teniendo en cuenta que es necesario calcular el sistema completo y luego una vez eliminando cada una de las sanguijuelas pegadas al parabrisas para saber si existe la posibilidad de salvarlo, hace que esta opción no sea la mas adecuada, sobre todo en aquellos problemas donde el parabrisas presenta una granularidad muy fina. Para esto, entonces, hacemos uso de una variación de la fórmula de Sherman–Morrison, evitando replantear todo el sistema desde cero, permitiendo reutilizar informacion de la ejecucion con todas las sanguijuelas y disminuyendo la complejidad a una mas adecuada.

Comencemos definiendo de manera mas adecuada la implementacion mas simple.

2.6 Implementacion Simple

Como ya hemos adelantado, la idea detras de este algoritmo sera quitar una sanguijuela del sistema, aplicarle eliminacion Gaussiana a la matriz obtenida y utilizar backwads substitution para obtener la solucion deseada.

Mas formalizado en pseudocodigo:

Luego, siendo la cantidad de sanguijuelas w , el tamaño de la matriz de resolución n y el tamaño de la banda p . La complejidad de este algoritmo sera $O(w * ((n * p^2) + (n * p)))$ que es lo mismo que $O(w * (n * p^2))$

Ahora intentaremos mejorar esta complejidad a travez de la utilizacion de la formula de sherman morrison.

TP1 3 void Ultimo_{disparo}()

- 1: Para cada sanguijuela
 - 2: La quito de la matriz A
 - 3: Quito su b en 0
 - 4: Aplico *Gauss*(A, b)
 - 5: Aplico *Backwards_ssubstitution*(A, b)
 - 6: Restablezco los valores originales en la matriz A y paso a la siguiente
-

2.7 Implementacion Sherman-Morrison

Entonces, sean A , A^* la matriz original del problema y la matriz modificada sin una de sus sanguijuelas respectivamente, queremos resolver $A^*x = b$ donde uv^t representa la matriz que introduce los cambios debido a la modificacion. Sea entonces $(A^*)x = b$. Resolvemos las siguientes ecuaciones, $Ay = b$ y $Az = u$ y obtenemos la nueva x a partir del siguiente cálculo:

$$x = y - \frac{zv^ty}{1 + v^tz}. \quad (5)$$

Así entonces, podemos obtener los resultados a partir de multiplicaciones de complejidad cuadrática y evitamos volver a usar el algoritmo de eliminación gaussiana de complejidad cúbica.

3 Discussion

3.1 Manipulación de estructura interna de la matriz

Al utilizar una estructura de menor espacio para representar la matriz del sistema, debimos lidiar con problemas de indexación. Gracias a una abstracción útil de la matriz bajo la clase *MatrizB*, entregando los métodos públicos *getVal* y *setVal*, que toman índices de la matriz original y los mapean a la representación interna que consta de un vector bidimensional, el problema pudo ser mitigado.

{Eficiencia temporal

La resolución de este sistema de ecuaciones implica una gran cantidad de operaciones aritméticas intervinientes. Trabajamos entonces en técnicas para disminuir el costo computacional sobre el algoritmo base de eliminación gaussiana a través de la explotación de características particulares del problema. Algunas técnicas, como la factorización LU, permiten una resolución más rápida del problema y funcionan para cualquier entrada. En cambio, otras técnicas aplican a un subconjunto de las posibles entradas.

Para la detección de sangijuelas que modifican levemente el sistema utilizamos la fórmula de Sherman-Morrison, obteniendo la variación del sistema de modo más eficiente que en el caso general que requiere de rehacer los cálculos de eliminación gaussiana.

Otra técnica que pensamos y finalmente no implementamos para detectar si eliminar una sangijuela salva el parabrisas en el caso en que exista una sangijuela cuyo radio de acción contiene al punto crítico es el siguiente:

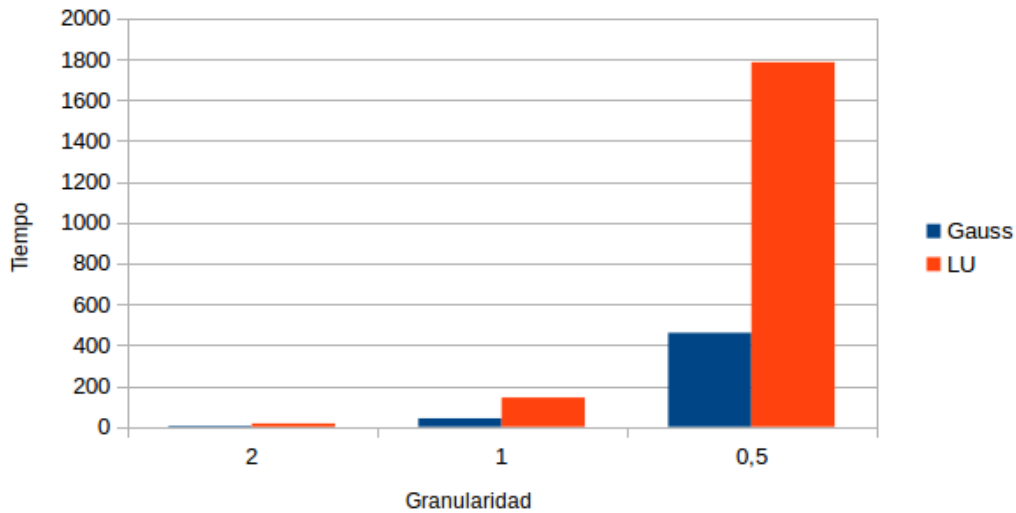
1. Si la sangijuela posee una temperatura menor a 235, eliminando cualquier otra el sistema va a seguir debajo de los 235 grados.
2. Si la sangijuela se encuentra a una temperatura igual o mayor a 235, probamos eliminarla.
3. Si eliminando la sangijuela se baja de los 235 grados, eliminandola logramos salvar el parabrisas, si no es cierto, no hay ninguna sangijuela que pueda ayudar ya que la sangijuela recién probada ejerce una temperatura directamente sobre el punto crítico.

4 Resultados

4.1 Análisis de temperatura y tiempos en función de las discretizaciones

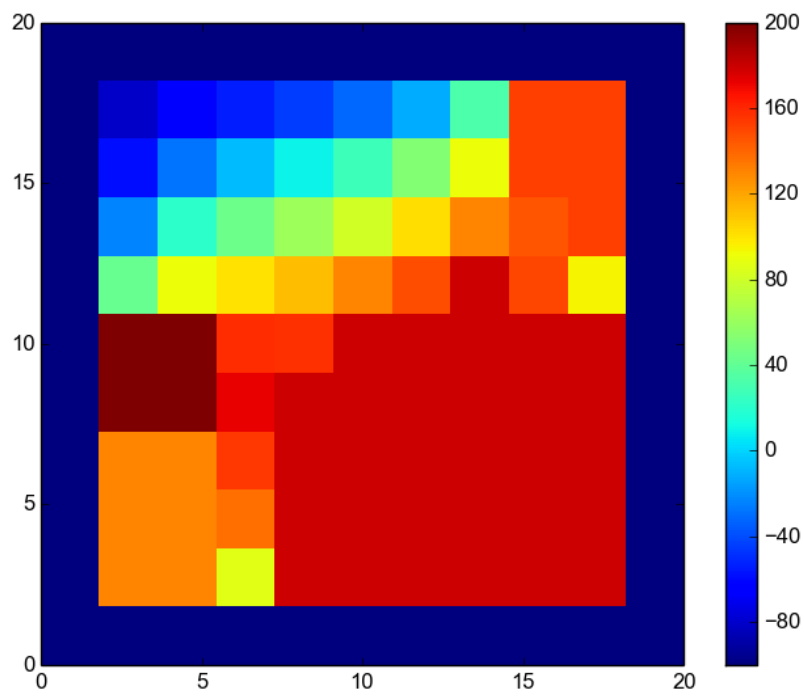
A continuación presentamos el análisis y las conclusiones obtenidas de acuerdo a la experimentación realizada. Como pudimos intuir en un primer momento, una mayor granularidad en la discretización hace escalar rápidamente la cantidad de ecuaciones y variables involucradas, con lo que el tiempo de los algoritmos escala rápidamente. Como supusimos, ambos algoritmos (Lease, la Eliminación Gaussiana explotando la estructura banda y la factorización LU) utilizados en el análisis, presentan diferencias de tiempo considerables, sobre todo a partir del último caso analizado donde se termina resolviendo un sistema de ecuaciones representado en una matriz de 200×200 . A continuación se presenta una lista con la granularidad utilizada y el tiempo obtenido para cada algoritmo:

Granularidad	Gauss	LU
2	4.678	15.877
1	41.381	143.102
0.5	459.756	1784.492
0.1	259967.136	1058285.72

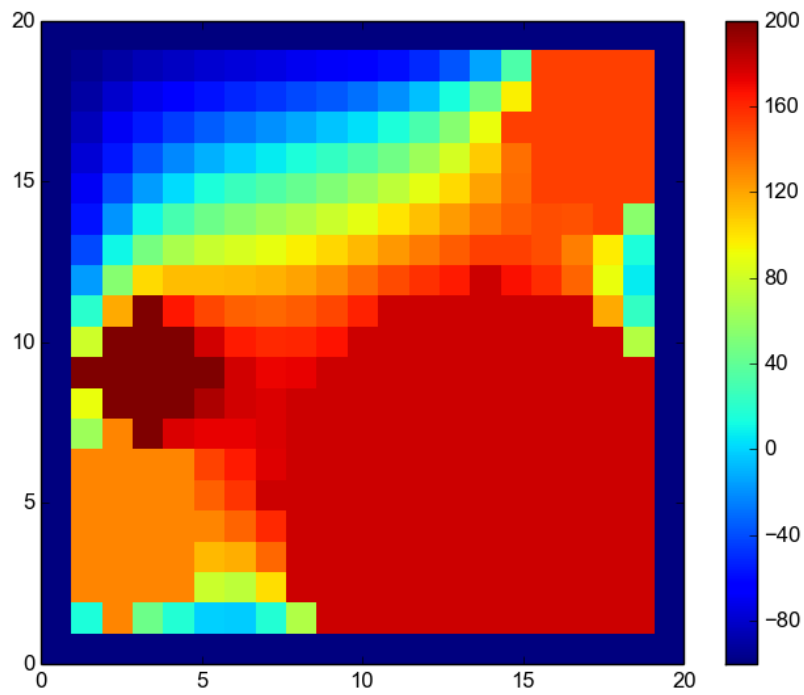


En cuanto a la distribución de calor, partiendo de la base de que un aumento en la granularidad permite una mejor representación de las sanguijuelas (son círculos), notamos que con este aumento y mejora en la representación, las temperaturas del punto crítico disminuyen considerablemente. Teniendo en cuenta esta información, queda claro como una baja granularidad impacta directamente sobre la precisión de los resultados (a expensas, como vimos con anterioridad, de los tiempos de cómputo).

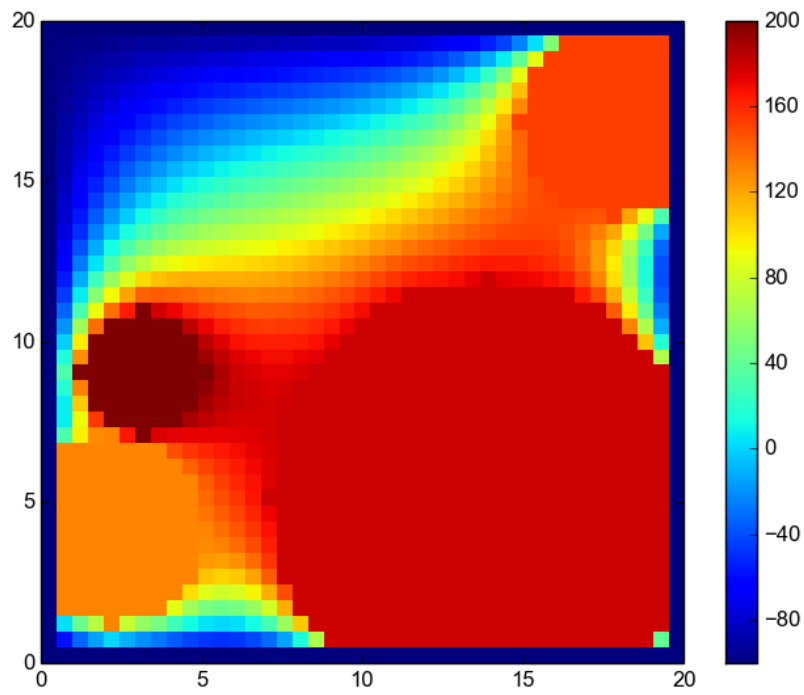
- Matriz 20×20 , granularidad 2.



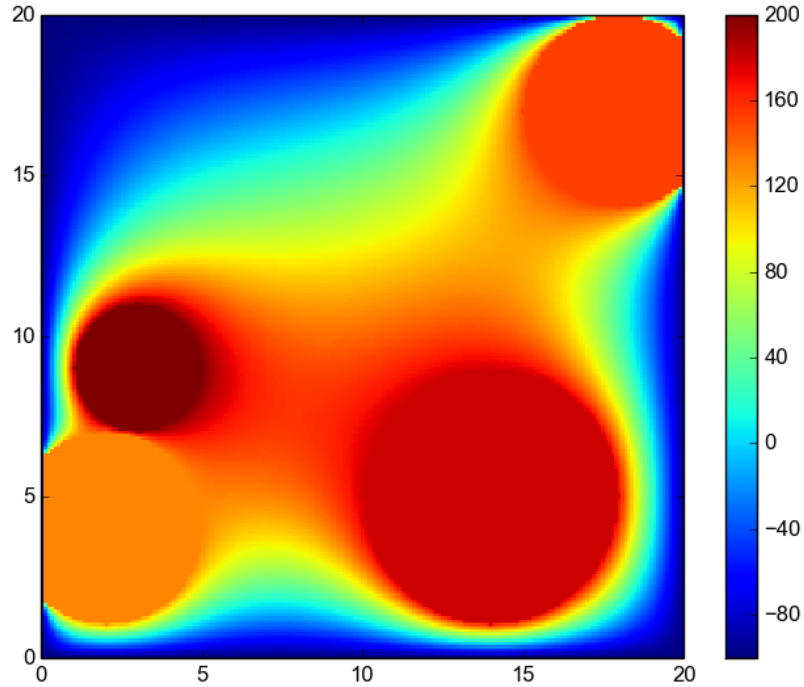
- Matriz 20×20 , granularidad 1.



- Matriz 20×20 , granularidad 0,5.

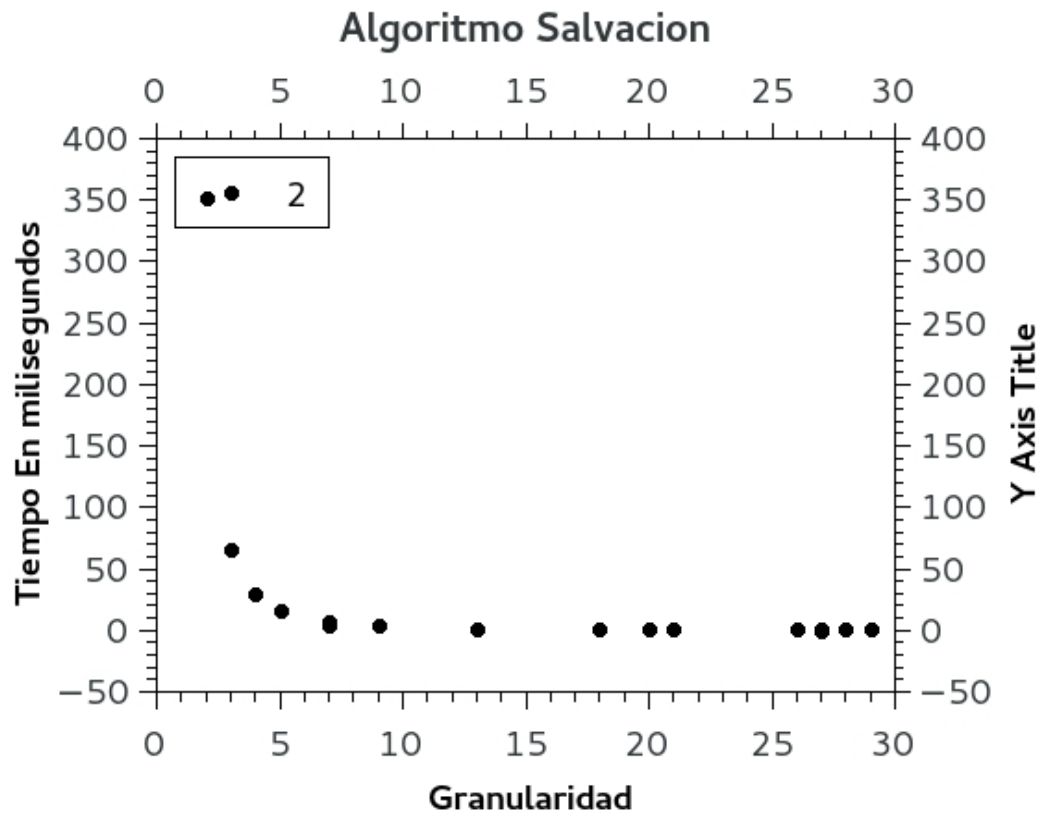


- Matriz 20×20 , granularidad 0,1.



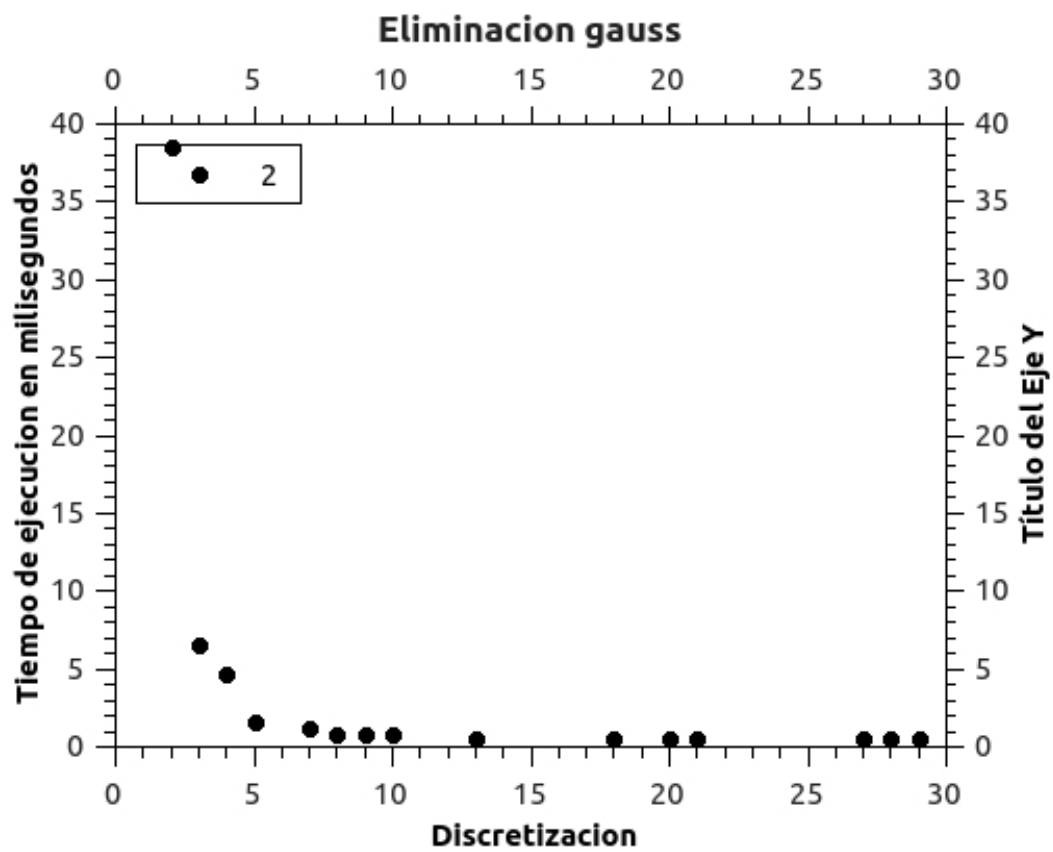
Ademas se hicieron test de stress con matrices de 50×50 generadas de forma aleatoria con una misma semilla y variando la granularidad para analizar la performances de los distintos algoritmos

Algoritmo de salvacion:

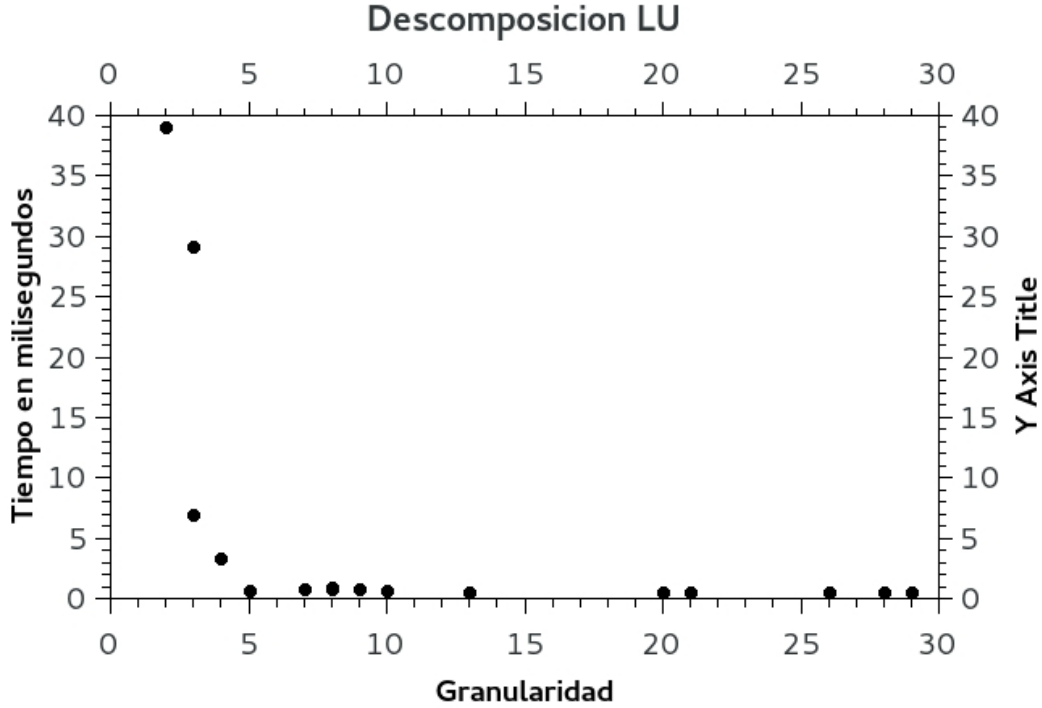


salvacion.jpg

Algoritmo de eliminacion gaussiana:



Algoritmo de LU:



4.2 Optimización del cómputo para cambios leves del sistema

Así como la particularidad de que el problema se puede plantear mediante el uso de una matriz banda permite ahorrar tiempo de cómputo y almacenamiento, otras especificidades del sistema o casos particulares permiten mejorar las estrategias de resolución. Para casos donde el sistema se modifica levemente, es decir que se produce un cambio en una sola fila de la matriz banda de resolución, es posible aplicar lo que se conoce como la fórmula de Sherman-Morrison. Lo que permite esta fórmula es evitar recomputar el algoritmo de eliminación gaussiana en estos casos, resumiendolo a una pequeña cantidad de operaciones aritméticas.

Para los casos donde aplica la fórmula de Sherman-Morrison, El cálculo sobre variación del sistema se realiza bajo una complejidad computacional de $O(n^2)$ mientras la resolución a través de la eliminación gaussiana clásica es de $O(n^3)$, notando una mejora sustancial usando Sherman-Morrison.

Si bien notamos esta mejora teórica y entendemos por qué sucede, no pudimos traducir esta mejora al código. Es decir, en los resultados obtenidos no se nota una mejoría respecto al caso base que es la aplicación de la Eliminación Gaussiana. Atribuimos esta diferencia entre los resultados teóricos y los observados a errores en el código y/o en las mediciones ya que no se presentan otros motivos para que los tiempos se mantengan más bajos utilizando Sherman-Morrison.

5 Conclusiones

5.1 Ventajas en el uso de una matriz banda

Una particularidad del sistema de ecuaciones planteado es que el valor de cada elemento del vector de resolución depende exclusivamente de, como máximo, otros 4 valores que además son cercanos al elemento a resolver, llamando cercano a un elemento que se encuentra a no más de una fila de distancia, dejando una gran cantidad de elementos (la gran mayoría en 0). Este problema en particular plantea una matriz banda con una cantidad de elementos no nulos mucho menor a aquellos que tienen un valor no nulo. Siendo más precisos la matriz tiene $(n \times m) \times (n \times m)$ elementos, con menos de $5 \times (m \times n)$ elementos no nulos. Gracias a eso es que pudimos realocar los valores en un espacio físico muy por debajo de lo que la representación matricial de un caso general requiere, mapeando los valores distintos de 0 a la matriz física y los valores nulos siendo entregados sin ser alojados en ninguna matriz.

5.2 Discretización de un problema continuo

Durante la resolución del problema pudimos ver que cambiando el h , es decir, la medida de discretización del continuo del parabrisas, obtenemos diferentes grados de error. A mayor h esperamos un error más pronunciado. Una posible definición de *Probabilidad* es el nivel de conocimiento sobre el sistema. Análogamente podemos decir que el nivel de discretización de un sistema continuo nos da una idea del conocimiento sobre el mismo.

Por otro lado, una mayor discretización del sistema impacta directamente sobre los tiempos de cómputo en su resolución. Queda entonces a criterio del desarrollador valorar ambos parámetros y balancear entre soluciones más precisas y tiempos de resolución más cortos. Esta elección está condicionada por los problemas acarreados provenientes de la resolución a través de computadoras, que no tienen la posibilidad de representar todos los números de la recta real, generando problemas de precisión.

5.3 Pérdida de precisión en el uso de aritmética de punto flotante

Así como la separación de los puntos del sistema queda, hasta cierto punto, a criterio de quien resuelve el problema, existe una gran limitación debido a la representación de los números de punto flotante en la computadora. Si bien C++ es independiente a la arquitectura, las arquitecturas modernas suelen utilizar el estándar IEEE754 para representar la recta numérica. Sabemos que realizando operaciones de punto flotante perdemos precisión, si bien esta puede ser calculada y acotada.

Entonces es así como la resolución del problema se ve afectado no solo por la reducción de un subespacio \mathbb{R}^2 en un conjunto finito de puntos, sino también por la imprecisión propia de la computadora al realizar los cálculos.

5.4 Otras aplicaciones

Vemos como este tipo de problema donde el valor de cada elemento de un espacio en 2 dimensiones (o más) depende del valor de elementos cercanos puede aplicarse en diversas áreas, por ejemplo en el procesamiento de imágenes, donde el suavizado en el zoom de una imagen se puede realizar a través del promedio de los vecinos de cada pixel. También es de posible aplicación para la propagación de ondas en distintos medios. Es interesante ver también como algunas técnicas genéricas ayudan a la velocidad de cómputo y luego se pueden realizar optimizaciones dentro del dominio de cada sistema en particular, como fue en nuestro caso Sherman-Morrison para casos donde el sistema varía levemente.

6 Apendice