

# Tarea 3: MÉTODO DE GRADIENTE CONJUGADO

Juan Gerardo Fuentes Almeida

**Abstract**—En esta práctica se implementa el algoritmo de Gradiente Conjugado en un problema de regularización de imágenes. Asimismo se hacen comparaciones en tiempo e iteraciones con el método de Gauss-Seidel.

## 1 INTRODUCCIÓN

Los métodos de Gradiente Conjugado se encuentran entre los más útiles para resolver sistemas grandes de ecuaciones, no requieren almacenamiento de matrices y nos permiten resolver problemas de optimización cuadrática.

La forma lineal del método de Gradiente Conjugado consiste en un método iterativo para resolver sistemas lineales con matrices definidas positivas. Su desempeño está determinado por la distribución de los eigenvalores de la matriz, esta distribución se puede transformar o pre-condicionar para mejorar significativamente la convergencia del método.

## 2 TEORÍA

### 2.1 Gradiente Conjugado Lineal

Se utiliza para resolver sistemas de ecuaciones lineales de la forma  $Ax = b$  para  $A$  simétrica positiva definida, se resuelve el problema de minimización

$$\min \phi(x) = \frac{1}{2}x^T Ax - b^T x$$

El gradiente  $\nabla\phi(x)$  equivale al residuo del sistema lineal

$$\nabla\phi(x) = Ax - b = r(x)$$

para una particular  $x = x_k$  se tiene  $r_k = Ax_k - b$

#### 2.1.1 Métodos de Dirección Conjugada

Los métodos de dirección conjugada se caracterizan por poder generar conjunto de vectores no nulos  $p_0, p_1, \dots, p_n$  tales que para una matriz simétrica y positiva definida  $A$  se tiene que

$$p_i^T A p_j = 0 \forall i \neq j$$

se dice entonces que estos vectores son *conjugados* respecto a la matriz  $A$

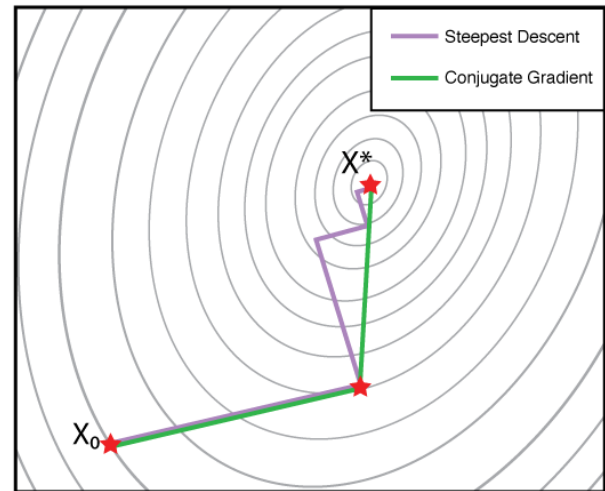
La importancia de esto es que podemos minimizar  $\phi$  en  $n$  pasos minimizando sucesivamente a lo largo de direcciones individuales en un conjunto conjugado. Dado un punto de inicio  $x_0$  y un conjunto de direcciones

conjugadas, podemos establecer la secuencia

$$x_{k+1} = x_k + \alpha_k p_k$$

donde  $\alpha_k$  es el minimizador de la función  $\phi$  determinada explícitamente por

$$\alpha_k = -\frac{g_k^T p_k}{p_k^T A p_k}$$



Dirección del Gradiente Conjugado

En el método de Gradiente Conjugado, cada dirección  $p_k$  se elige como una combinación lineal del residuo negativo  $-r_k$ , el cual es la dirección de máximo descenso para la función  $\phi$ , y la dirección anterior  $p_{k-1}$

$$p_k = -g_k + \beta_k p_{k-1}$$

donde el escalar  $\beta_k$  se determina de forma que  $p_k$  y  $p_{k-1}$  sean conjugados respecto a  $A$ :

$$\beta_k = \frac{g_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$$

En esta práctica calculamos  $\beta_k$  de acuerdo a la definición de Fletcher-Reeves:

$$\beta_k = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}$$

### 3 RESULTADOS

Como parte de la práctica, se tomó un conjunto de imágenes de resonancia magnética para implementar un registro rígido y minimizar la función de error determinada por la siguiente expresión:

$$\min_{\alpha} f(\alpha) = \frac{1}{2} \sum_{x \in \Omega} \left( I_2(x) - I_1(x + X\alpha) - \nabla I_1(x + X\alpha)^T X\alpha \right)^2$$

$I_2$  es una imagen de referencia,  $I_1$  es la imagen observada cuyo dominio está definido en  $x = x_1, x_2$ , de manera que  $\omega$  representa el conjunto de todos los píxeles de la imagen,  $X$  y  $\alpha$  son variables definidas de la siguiente manera:

$$X = \begin{bmatrix} x_1 & x_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & 1 \end{bmatrix} \text{ y } \alpha = \begin{bmatrix} \cos(\theta) - 1 \\ \sin(\theta) \\ v_1 \\ -\sin(\theta) \\ \cos(\theta) - 1 \\ v_2 \end{bmatrix}$$

Esto permite que la transformación de una imagen a otra este dada por el producto  $x + X\alpha$

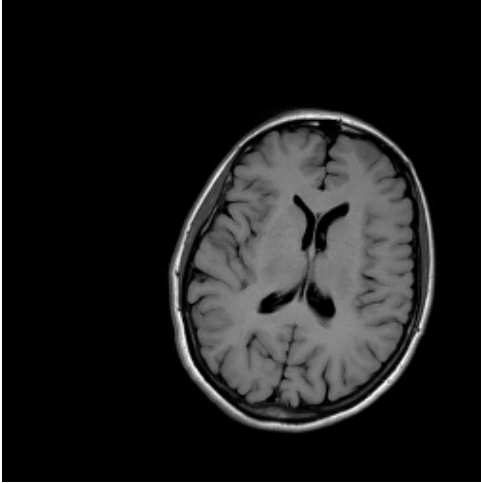


Imagen Inicial u Observada

El gradiente y el Hessiano de esta función están definidos de la siguiente manera:

$$\nabla f(\alpha) = \sum [I_2(x) - I_1(x)] [-X^T \nabla I_1(x + X\alpha)]$$

$$\nabla^2 f(\alpha) = \sum [X^T \nabla I_1(x + X\alpha)] [X^T \nabla I_1(x + X\alpha)]^T$$

En la práctica, estas funciones fueron las que nos dieron mejores resultados.

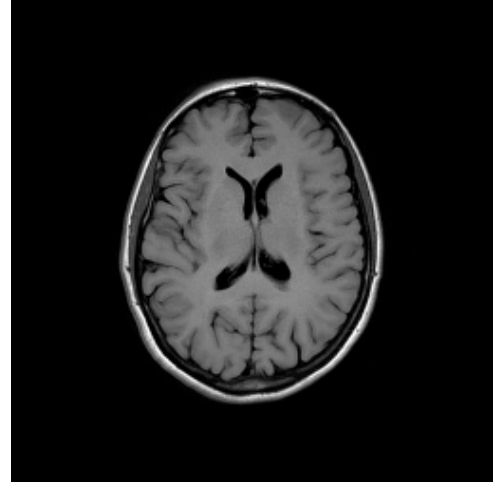


Imagen de referencia

El algoritmo lee los archivos de las dos imágenes y realiza la minimización de la función objetivo con el método de la Región de Confianza, en conjunto con el método de Dogleg para el cálculo del punto de Cauchy; se inicia con un desplazamiento inicial de 10 píxeles hacia arriba y hacia la izquierda para acercarnos un poco a la posición de la imagen de referencia.

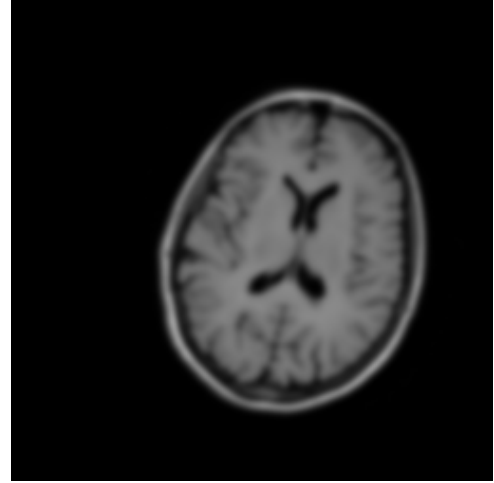


Imagen inicial, se tuvo que hacer un emborronado de la imagen para obtener mejores resultados

Aunque la función objetivo siempre muestra valores demasiado grandes, se pudo observar una mejoría en la posición y orientación de la imagen con respecto a la de referencia. Otro comportamiento observado fue que a partir de cierto número de iteraciones, la función objetivo deja de disminuir quedándose en un valor muy grande y reduciendo prácticamente a cero la región de confianza.

### 4 CONCLUSIONES

En general no se obtuvieron los resultados que se esperaban, se tuvieron muchos problemas, principalmente con respecto a la correcta definición de la función objetivo y sus derivadas, también se tuvieron problemas con los métodos

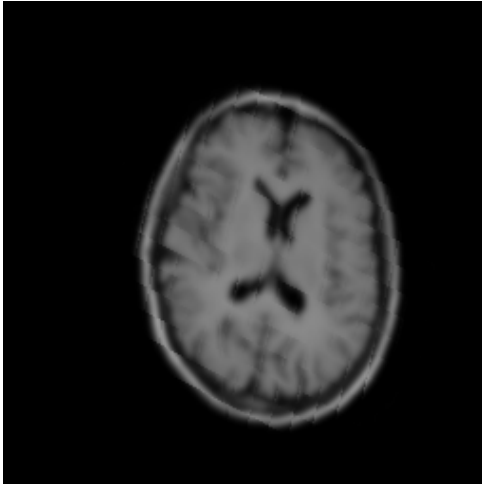


Imagen final despues de 20 mejoras, se aprecia la rotacion que se produce para empalmarse con la imagen de referencia

de optimización vistos en clase, los cuales directamente aplicados a este problema no producían los resultados esperados. Se espera que en futuras prácticas se afinen más detalles sobre este tipo de implementaciones.

## REFERENCES

- [1] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer series in operations research and financial engineering. Springer, New York, NY, 2. ed. edition, 2006.

## APPENDIX

El programa está implementado tomando en cuenta todas estandarizaciones indicadas en el curso.

Un *makefile* ha sido generado, el cual soporta los comandos *make*, *run* and *clean*. El programa recibe como primer argumento el nombre del archivo de la imagen inicial, y como segundo argumento el nombre del archivo de la imagen de referencia.