

OPTIMIZATION. HOMEWORK 10

OSCAR DALMAU

Conjunto de Problemas

- (1) Implementar el método del gradiente conjugado No-lineal con las cuatro opciones de calcular β vistas en clase: Fletcher-Reeves (FR), Polak-Ribiere (PR), Hestenes-Stiefel (HS) y la que combina FR con PR.
- (2) Con la implementación anterior, resolver el siguiente problema de optimización no lineal:

$$(1) \quad \min_X \sum_{i,j} \left[(x_{i,j} - g_{i,j})^2 + \lambda \sum_{(l,m) \in \Omega_{i,j}} \sqrt{(x_{i,j} - x_{l,m})^2 + \mu} \right]$$

donde $\mu = 0.01$, $\lambda > 0$ es un parámetro de regularización dado; g es la función a la cual se le desea filtrar el ruido, por ejemplo una imagen; $\Omega_{i,j}$ es el conjunto de índices formados por los vecinos del píxel (i, j) , es decir $\Omega_{i,j} = \{(i+1, j), (i-1, j), (i, j+1), (i, j-1)\}$. Observe que si la imagen original es de tamaño $m \times n$ entonces la imagen con el ruido filtrado será $X^* = [x_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$ donde X^* es el minimizador del problema.

Especificaciones

En total deberán resolver el problema (1) para las cuatro distintas opciones de β . Para cada β (es decir, calculadas mediante Fletcher-Reeves (FR), Polak-Ribiere (PR), Hestenes-Stiefel (HS) y la que combina FR con PR), reportar la gráfica de las iteraciones vs la función objetivo, también deberán reportar la gráfica de las iteraciones vs la norma del gradiente. Además deberán reportar la imagen original y la imagen con el ruido filtrado es decir la imagen $X^* = [x_{i,j}]_{i,j}$.

También deberán reportar una pequeña tabla comparativa, donde se comparen el número de iteraciones y el tiempo que tarda el método en resolver el problema para cada parámetro β .

La función **main** deberá recibir los siguientes parámetros:

- (1) X_0 : el punto inicial.
- (2) tol_g : la tolerancia para la norma del gradiente.

- (3) f : la función objetivo
- (4) $grad$: el gradiente
- (5) g : la función g del modelo de optimización (1), por ejemplo una imagen.
- (6) $beta$: un string con las opciones: 'PR' (Polak-Ribiere), 'FR' (Fletcher-Reeves), 'HS' (Hestenes-Stiefel) y 'FRPR' (la mezcla de Fletcher-Reeves y Polak-Ribiere).

Observe que para garantizar que las direcciones de búsquedas empleadas por el “algoritmo del gradiente conjugado no lineal” sean direcciones de descenso, se debe realizar búsqueda en línea para calcular el tamaño de paso (para ello puede usar las funciones de `scipy`).

Como criterio de parada de su algoritmo deben usar el criterio de paro de la norma del gradiente y un número máximo de iteraciones. La imagen g corresponde a la imagen de “lena”.