Tarea final 2023

14 de noviembre de 2023

Ejercicio 1 - Puntos de Fekete Logarítmico (40)

Un problema muy importante consiste en distribuir n puntos $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ sobre la superficie de la esfera $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$ de manera de minimizar una cierta energía E. Es un problema fundamental resuelto para ciertas configuraciones de n y ciertas funciones de energía, y abierto para otras. En este ejercicio vamos a buscar un mínimo local de:

$$E(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = -\sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \log(\|\mathbf{x}_i - \mathbf{x}_j\|) \quad \mathbf{x}_k \in \mathcal{S} \; i = 1, \ldots, n.$$

a) Calcular el gradiente de E respecto a un punto $\mathbf{x}_i$, $\nabla_i E(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in \mathbb{R}^3$.

b) Implementar y probar el siguiente método: en la iteración k se actualiza el punto $\mathbf{x}_i$ donde $i = k \bmod n$ (o sea que $i = 0, 1, \ldots, n-2, n-1, 0, 1, \ldots, n-2, n-1, 0, \ldots$). La actualización en cuestión consiste en ejecutar una iteración de gradiente proyectado con *paso* $\alpha > 0$ donde el conjunto factible es $\mathcal{S}$ y la dirección del gradiente es respecto al punto $\mathbf{x}_i$. El método se detiene cuando la variación de $\mathbf{x}_i$ tenga norma menor a cierto valor ϵ.

c) Probar el método anterior para $n = 100$, $\alpha = 10^{-2}$ y $\epsilon = 10^{-3}$. Los datos iniciales se deben generar usando la función `inicializar_fekete` del archivo `util.py`

d) Graficar la energía en función de las iteraciones y los puntos al inicio y al final de la ejecución usando la función `mostrar_fekete` del archivo `util.py`. Comente sus resultados.

Ejercicio 2 - Otra vez LASSO (60)

En el Obligatorio 4 vimos cómo resolver el siguiente problema de regresión lineal regularizada, conocido como LASSO, utilizando métodos proximales y el ADMM:

$$(\mathbf{P0})\ \mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \theta\|\mathbf{x}\|_1,$$

en donde $\mathbf{A} \in \mathbb{R}^{m\times p}$ e $\mathbf{y} \in \mathbb{R}^p$ son datos del problema y θ es un parámetro de regularización. La idea de este ejercicio es resolverlo utilizando otros métodos y comparar sus desempeños.

Datos y parámetros del problema : Los datos para este ejercicio son generados por las funciones `generar_A` y `generar_y` del archivo `util.py` provisto junto con la letra. Fijamos $\theta = 0{,}1$. La condición de parada para los métodos iterativos a implementar será que la norma del iterando $\mathbf{x}$ no tenga una variación relativa mayor a 10^{-5} en la iteración k:

$$\frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|}{\|\mathbf{x}^k\|} \le 10^{-5}$$

a) Plantear el LASSO como un problema de optimización cuadrático con restricciones lineales (al que llamaremos P1). Sugerencia: el problema $\min_x |f(x)|$ se puede sustituir por

$$\begin{aligned}
&\min_t\ t \\
&\text{s.a.} \\
&\qquad t \ge f(x) \\
&\qquad t \ge -f(x)
\end{aligned}$$

b) Resolver el problema (P1) utilizando el paquete CVXPY (https://www.cvxpy.org/).

c) Resolver el problema original (P0) utilizando el método de *descenso por coordenadas*. Este método consiste en actualizar de a una coordenada de la solución por vez, dejando el resto fijas:

$$x_i^k = \arg\min_{\zeta} \frac{1}{2}\|\mathbf{A}(x_1^{k-1},\ldots,x_{i-1}^{k-1},\zeta,x_{i+1}^{k-1},\ldots,x_p^{k-1}) - \mathbf{y}\|_2^2 + \theta\|(x_1^{k-1},\ldots,x_{i-1}^{k-1},\zeta,x_{i+1}^{k-1},\ldots,x_p^{k-1})\|_1,$$

en donde $i = k \bmod p$ ($i = 0,1,\ldots,k-2,k-1,0,1,\ldots,k-2,k-1,0,\ldots$). Nota: este método puede requerir *muchas* iteraciones, pero la idea es ahorrar cómputo de manera de que cada iteración sea muy rápida. *intente identificar cálculos repetitivos y constantes dentro de las iteraciones para acelerar el proceso. Puede servir lo siguiente:*

$$\mathbf{A}\mathbf{x} = \sum_{j\neq i} \mathbf{A}_j x_j + \mathbf{A}_i x_i,$$

donde $\mathbf{A}_i$ es la columna i-ésima de $\mathbf{A}$.

d) Resolver el problema (P0) utilizando el ADMM implementado en el Obligatorio 4; pruebe con distintos valores del parámetro λ del ADMM (se sugiere probar potencias de 10).

e) Compare los tiempos de ejecución, las soluciones y la función de costo obtenida. Comente sobre las diferencias numéricas que observa entre las distintas soluciones.