Tonantzin Real Rojas
163491

1 Examen parcial

1.1 Teoría

1.1.1 ELI5 (30 puntos)

Explícame como si tuviera 5 años. Es un subreddit muy popular donde se explican conceptos muy interesantes de maneras muy sencillas. Richard Feynman tenía una metodologia de aprendizaje que involucra explicar conceptos complejos a un niño. Por ende para esta parte del examen tenemos lo siguiente:

- Explica el algoritmo de búsqueda lineal a un niño de 5 años.
- Explica el algoritmo de región de confianza a un niño de 5 años.

1.1.2 Demostración (30 puntos)

Si tenemos f una cuadrática convexa $f(x) = \frac{1}{2}x^TQx - b^Tx$. Demuestra que el minimizador de una dimensión sobre la linea $x_k + \alpha p_k$ es:

$$\alpha_k = -\frac{\nabla f_k^T p_k}{p_k^T Q p_k}$$

1.2 Code (40 puntos)

En la carpeta data hay una base de datos. Que contiene datos de delitos cometidos en la Ciudad de México, para mayo referencia de cómo obtener estos datos, ver esta página.

La jefa de gobierno te dice que hubo un problema y que vamos a reinstalar las cámaras de seguridad de los sectores que componen esta base de datos, esto implica que podemos ubicar las cámaras en mejores lugares a los que están. No tenemos certeza de cuántas cámaras serán, pero para simplicidad del problema asumamos que tenemos 8000 cámaras de seguridad. Te piden ayuda para elegir la posición óptima de las 8000 nuevas cámaras.

Establece claramente la métrica que vas a buscar optimizar y cuál es el razonamiento para resolver el problema de colocar las cámaras asi como generar una función que calcule, para un conjunto de 8000 puntos, la función de costo que se optimizará.

1.1.1 ELI5 (30 puntos)

Explícame como si tuviera 5 años. Es un subreddit muy popular donde se explican conceptos muy interesantes de maneras muy sencillas. Richard Feynman tenía una metodologia de aprendizaje que involucra explicar conceptos complejos a un niño. Por ende para esta parte del examen tenemos lo siguiente:

- Explica el algoritmo de búsqueda lineal a un niño de 5 años.
- Explica el algoritmo de región de confianza a un niño de 5 años.

Búsqueda lineal

El algoritmo elige una dirección p_k y busca a lo largo de ella a partir de la iteracción actual x_k una nueva iteración con un valor de función más bajo

Para saber cuánto nos movemos sobre p_k debemos encontrar al tamaño de paso α y se hace mediante el problema de minimización:

$$\min_{\alpha>0} f(x_k+\alpha p_k) \qquad (1)$$

Este método genera cierto número de pasos para probar hasta que encuentra a uno que se aproxima al mínimo de (1)

En cada nuevo punto, se buscan nuevas direcciones de búsqueda y de tamaños de pasos y se repite el proceso

→ se fija una dirección y se elige el tamaño de paso

En otras palabras: la BL encuentra una direccion de descenso en la que la función f toma un valor más pequeño; es decir, en la que f se reduce y luego calcula el tamaño de paso que determina cuánto se debe mover x en esa dirección

Región de confianza

La información recabada de f se usa para construir una función modelo m_k cuyo comportamiento cerca del punto actual x_k es similar al de la función objetivo acrual f. Como el modelo m_k puede no ser una buena aproximación de f cuando x está lejos de x_k, restringimos la búsqueda de un minimizador de m_k a una región alrededor de x_k.

Encontramos el paso candidato p mediante resolver aproximadamente el subproblema:

$$\min_{p} m_k(x_k+p) \qquad \text{donde } x_k+p \text{ está dentro de la región de confianza}$$

Si la solución candidata no produce un descenso suficiente en f, se concluye que la región de confianza es muy larga y la encogemos y repetimos

Usualmente la región de confianza es una bola definida por $\|p\|_2 < \Delta$ donde el escalar $\Delta > 0$ es el radio de la región de confianza

Aunque no necesariamente es una bola

El modelo m_k usualmente se define como

$$m_k(x_k + p) = f_k + p^T \nabla f_k + \frac{1}{2} p^T B_k p$$

donde $f_k, \nabla f_k, B_k$ son escalar, vector y matriz respectivamente

↳ Hessiana $\nabla^2 f_k$ o una aproximación a ella

En otras palabras: la región de confianza es una parte de la región de la función objetivo. Esta región (la de confianza) se determina mediante una aproximación que resulta de una función como auxiliar m_k que cerca del punto de interés x_k tiene un comportamiento similar al de la función f original

Puede que en ocasiones el tamaño de esta región de confianza no sea el correcto, en cuyo caso se ajusta su tamaño y se repite el proceso antes mencionado

En resumen:

Búsqueda lineal

a) Fija la dirección p_k

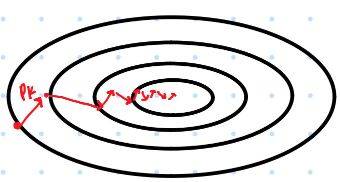
b) Identifica una distancia apropiada α_k (tamaño de paso)

Región de confianza

a) Elegir distancia máxima (radio de la región) Δ_k

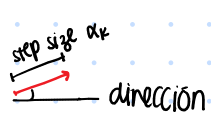
b) Buscamos una dirección y paso que mejoren más la aproximación dada la restricción de la distancia

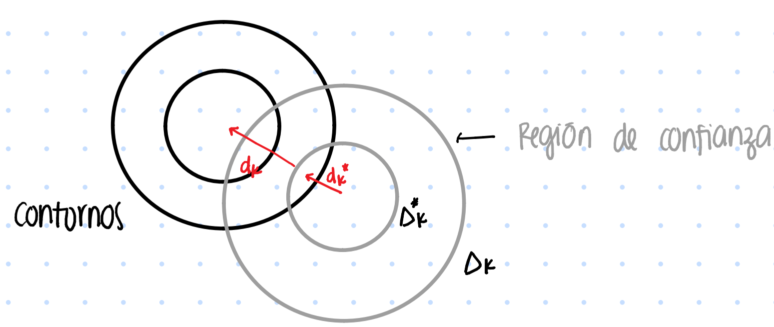
c) Si el paso no satisface, se reduce la distancia Δ_k y repetimos



Line search

con





Región de confianza

1.1.2 Demostración (30 puntos)

Si tenemos f una cuadrática convexa $f(x) = \frac{1}{2}x^TQx - b^Tx$. Demuestra que el minimizador de una dimensión sobre la linea $x_k + \alpha p_k$ es:

$$\alpha_k = -\frac{\nabla f_k^T p_k}{p_k^T Q p_k}$$

Consideremos la f. cuadrática convexa fuerte

$$f(x) = \frac{1}{2}x^TQx - b^Tx$$

Buscamos a un minimizador en el rayo $x_k + \alpha p_k$

i.e. $\frac{d}{d\alpha}(f(x_k + \alpha p_k)) = 0$

Entonces

$$f(x_k+\alpha p_k) = \frac{1}{2}(x_k+\alpha p_k)^T Q(x_k+\alpha p_k) - b^T(x_k+\alpha p_k)$$

$$= \frac{1}{2}x_k^TQx_k + \frac{1}{2}x_k^TQ\alpha p_k + \frac{1}{2}\alpha p_k^TQx_k + \frac{1}{2}\alpha^2 p_k^TQp_k - b^Tx_k - \alpha b^Tp_k$$

$$= \frac{1}{2}x_k^TQx_k + \alpha x_k^TQp_k + \frac{1}{2}\alpha^2 p_k^TQp_k - b^Tx_k - \alpha b^Tp_k$$

$$\Rightarrow \frac{d}{d\alpha}(f(x_k+\alpha p_k)) = x_k^TQp_k + \alpha p_k^TQp_k - b^Tp_k$$

$$= \alpha p_k^TQp_k + (x_k^TQ - b^T)p_k$$

$$= \alpha p_k^TQp_k + \nabla f_k^T p_k$$

$\nabla f_k = Qx_k - b$

Por lo que para el mínimo necesitamos $\frac{d}{d\alpha}(f(x_k+\alpha p_k)) = 0$

$$\Rightarrow \alpha p_k^TQp_k = -\nabla f_k^T p_k$$

$$\therefore \alpha = \frac{-\nabla f_k^T p_k}{p_k^TQp_k}$$

$\blacksquare$

1.2 Code (40 puntos)

En la carpeta data hay una base de datos. Que contiene datos de delitos cometidos en la Ciudad de México, para mayo referencia de cómo obtener estos datos, ver esta página.

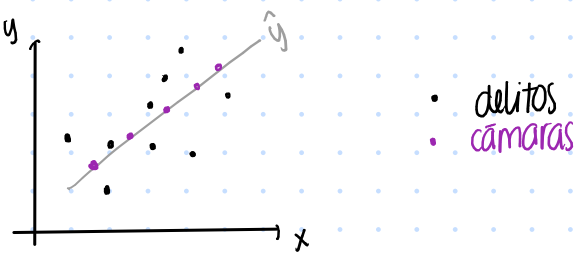
La jefa de gobierno te dice que hubo un problema y que vamos a reinstalar las cámaras de seguridad de los sectores que componen esta base de datos, esto implica que podemos ubicar las cámaras en mejores lugares a los que están. No tenemos certeza de cuántas cámaras serán, pero para simplicidad del problema asumamos que tenemos 8000 cámaras de seguridad. Te piden ayuda para elegir la posición óptima de las 8000 nuevas cámaras.

Establece claramente la métrica que vas a buscar optimizar y cuál es el razonamiento para resolver el problema de colocar las cámaras asi como generar una función que calcule, para un conjunto de 8000 puntos, la función de costo que se optimizará.

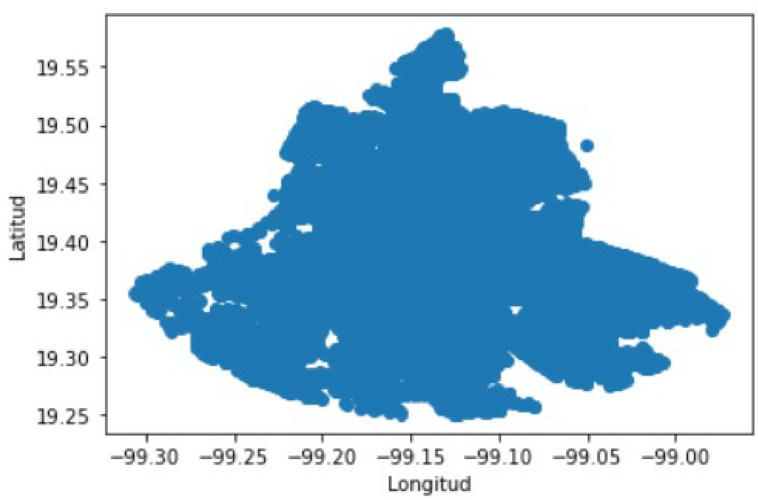
Conocemos las latitudes (x) y longitudes (y) donde sucedieron los delitos lo que queremos hacer es determinar en qué lugar vamos a poner las 8000 cámaras

una idea para determinar la posición de las cámaras es ubicarlas en los lugares donde sucedan más crímines. Para hacerlo podemos proponer una función f (puede ser polinomial) tal que la distancia entre la cámara y los delitos sea mínima (la idea algo similar a mínimos cuadráticos)

Buscamos



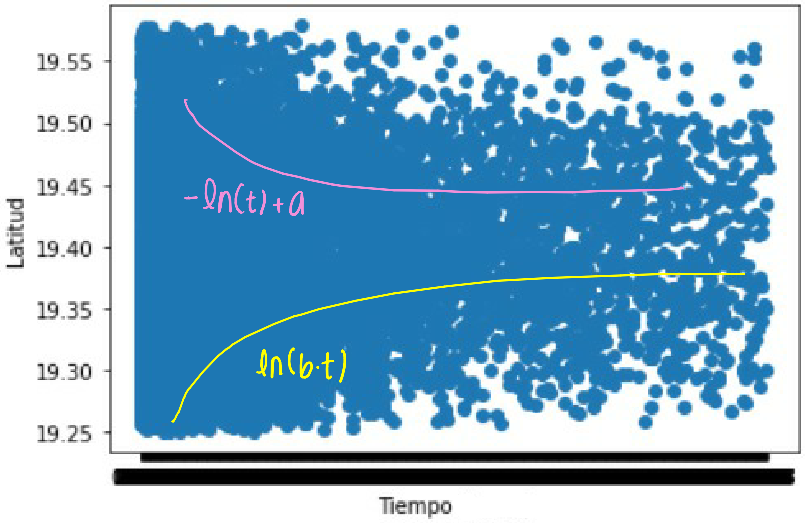
Para saber qué función proponer primero exploremos cómo se ven nuestros datos de delitos



Entonces lo que queremos hacer es dada la figura de arriba, poner 8000 cámaras en los lugares tales que estas estén en los lugares donde suceden más delitos

Tenemos que cada y_i se tomó en un tiempo t_i

Al comparar las y_i contra el tiempo t_i obtenemos



El eje x se refiere al tiempo pero como está en formato h:m:s por eso se ve muy amontonado (no sé cómo cambiar eso)

Notemos que si elegimos $\ln(bt)$ y $-\ln(t)+a$ entonces parece que se acopla relativamente bien a nuestros datos

$\Rightarrow$ una función que podemos proponer es

$$\phi(t,x) = x_1 + x_1\ln(t) - \ln(x_2\cdot t)$$

Con lo anterior tenemos al residual

$$r_j(x) = y_j - \phi(t_j, x) \qquad \text{con } j = 1, \dots, m \quad \text{y} \quad m = 8000$$

que corresponde al # de cámaras que vamos a poner

De esta manera la función de costo a optimizar es

$$f(x) = \sum_{i=1}^{m} r_i(x)^2$$

Y la función objetivo es

$$\min_x f(x)$$

El resto del problema está en Python en el archivo 1.2 Delitos CDMX