(1) Tenemos un problema. Hay montañas, valles y planicies y queremos encontrar el hoyo o cráter más profundo que haya. Lo buscaremos paso por paso hasta que no se pueda progesar más o hasta que nuestro cráter sea lo suficientemente profundo. ¿Cómo lo logramos?

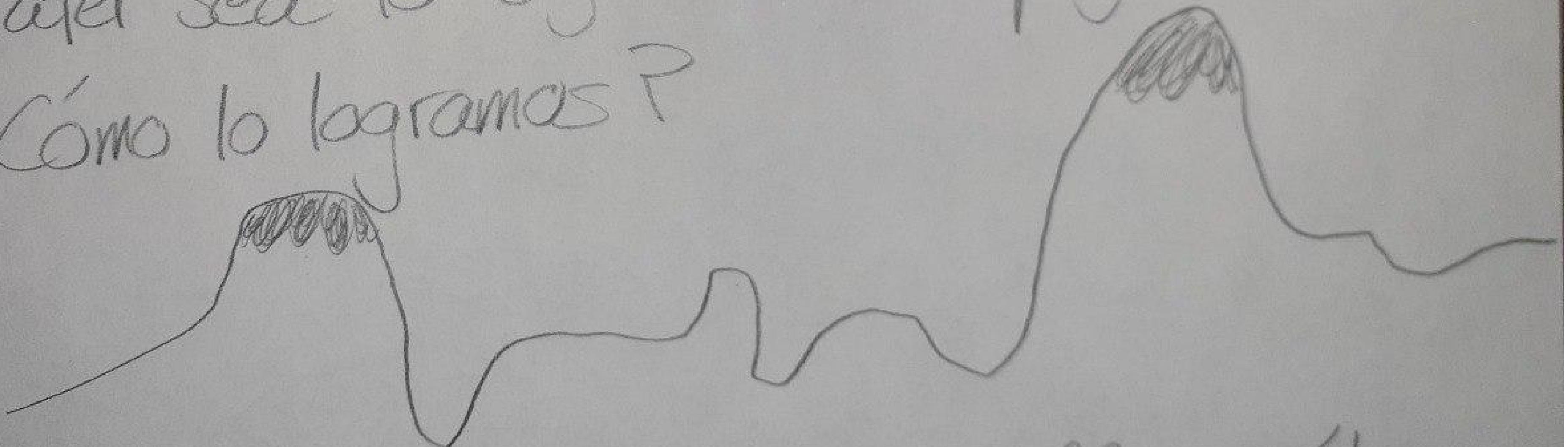

Figura 1. Montañas planicies y cráteres

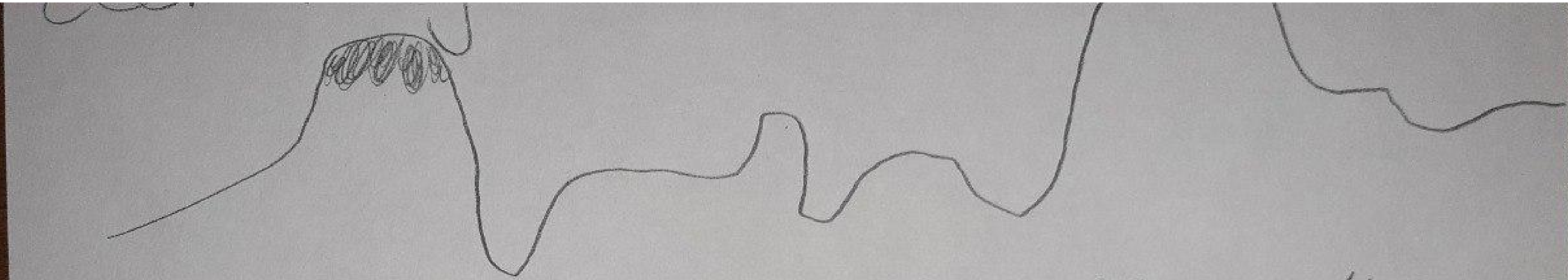

Figura 1 Montañas planicies y cráteres

Tenemos que trabajar a partir de la información de nuestra altura y nuestros pasos.

¿Cómo elegimos los pasos?

ⓐ Búsqueda lineal

La primera idea es elegir una dirección para desplazarnos y dar pasos en tal dirección para probar si nuestra altitud disminuye. Para elegir la distancia de nuestros pasos utilizaremos un ninyutsu matemático superpoderoso:

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

Con esta técnica ninya hallaremos la longitud del paso. Nuestras decisiones las tomaremos con α, alfa, la longitud del paso, x_k, nuestro paso actual,

p_k: nuestra dirección

del paso. Nuestras decisiones las tomaremos con α, alfa, la longitud del paso, x_k, nuestro paso actual,

p_k: nuestra dirección

f: nuestro medidor de altitud.

Con nuestro amigo min buscaremos el mínimo.

Paso tras paso actualizaremos la ubicación de nuestro paso. Nos detendremos cuando hayamos cumplido con las indicaciones del señor Wolfe o del señor Goldstein. Nos indican que debemos bajar lo suficiente con nuestros pasos

• y pasos demasiado pequeños no son aceptables.
Caminamos mentalmente viendo el camino hasta que al cumplir estas condiciones veamos cual fue la distancia de nuestros pasos. Tomamos esta información "del futuro" y con ella damos nuestros pasos. Esto es la retropropagación.

Hay otro método.
(b) Región de confianza.
¿Y si construimos un modelo similar al paisaje?

nuestro paso.

Hay otro método.

(b) Región de confianza.

¿Y si construimos un modelo similar al paisaje segmento por segmento? Hay que tener cuidado porque podría alejarse mucho de la realidad.

Esta vez empezaremos por elegir una distancia máxima para hacer nuestro modelo imaginario del paisaje y después elegiremos nuestra dirección y la longitud de nuestros pasos.

Si nuestros pasos non son triunfadores reducimos el tamaño de nuestro modelo imaginario. Repetimos hasta que estamos lo suficientemente cerca del objetivo o hasta que no podamos progresar más.

1.2

$$\phi(\alpha_k) = f(x_k + \alpha_k p_k)$$

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

$$= \frac{1}{2} x_k^T Q x_k + \frac{1}{2}\alpha_k^2 p_k^T Q p_k + \frac{1}{2} x_k^T Q p_k \alpha_k$$

$$+ \alpha_k \frac{1}{2} p_k^T Q x_k - b^T x_k - \alpha_k b^T p_k$$

$$\therefore \phi'(\alpha_k) = \alpha_k p_k^T Q p_k + \frac{1}{2} x_k^T Q p_k + \frac{1}{2} p_k^T Q x_k$$

$$- b^T p_k$$

Tenemos que $b = \nabla f_k - Q x_k$

$$\Rightarrow \phi'(\alpha_k) = \alpha_k p_k^T Q p_k + \frac{1}{2} x_k^T Q p_k + \frac{1}{2} p_k^T Q x_k$$

$$- (Q x_k - \nabla f_k)^T p_k$$

$$\therefore \phi'(\alpha_k) = \alpha_k p_k^T Q p_k + \frac{1}{2} x_k^T Q p_k + \frac{1}{2} p_k^T Q x_k - b^T p_k$$

Tenemos que $b = \nabla f_k - Q x_k$

$$\Rightarrow \phi'(\alpha_k) = \alpha_k p_k^T Q p_k + \frac{1}{2} x_k^T Q p_k + \frac{1}{2} p_k^T Q x_k - (Q x_k - \nabla f_k)^T p_k$$

$$= \alpha_k p_k^T Q p_k + \cancel{\frac{1}{2} x_k^T Q p_k} + \cancel{\frac{1}{2} p_k^T Q x_k} - \cancel{x_k^T Q p_k} + \nabla f_k^T p_k$$

Entonces $\phi'(\alpha_k) = 0 \iff 0 = \alpha_k p_k^T Q p_k + \nabla f_k^T p_k$

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

Con el algoritmo de las k-medias, en este caso 8000 medias, podemos resolver el problema. (2)

Dadas las n coordenadadas de los crímenes: $\{(x_1, y_1), \ldots, (x_n, y_n)\} = \mathcal{X}$, buscamos agruparlas en 8000 conjuntos $(n \geq 8000)$, es decir la cámara más cercana, $\mathcal{S} = \{S_1, \ldots, S_{8000}\}$ para minimizar la suma de cuadrados dentro del respectivo grupo (cámara) más cercana a un crimen.

$$\underset{S}{\operatorname{argmin}} \sum_{i=1}^{8000} \sum_{x \in S_i} \|x - \mu_i\|_2^2 = \underset{S}{\operatorname{argmin}} \sum_{i=1}^{8000} |S_i| \operatorname{Var}(S_i)$$

8000 conjuntos ($n \geq 8000$), es decir la cámara más cercana, $S = \{S_1, \ldots, S_{8000}\}$ para minimizar la suma de cuadrados dentro del respectivo grupo (cámara) más cercana a un crimen.

$$\underset{S}{\operatorname{argmin}} \sum_{i=1}^{8000} \sum_{x \in S_i} \|x - \mu_i\|_2^2 = \underset{S}{\operatorname{argmin}} \sum_{i=1}^{8000} |S_i| \operatorname{Var}(S_i)$$

con $x \in X$

donde μ_i es la media de puntos en S_i.

μ_i sería la ubicación de la cámara y S_i el grupo de los crímenes cercanos a μ_i.

Procedimiento

Dado un conjunto de 8000 medias $m_1^{(1)}, \ldots, m_{8000}^{(1)}$

alternaremos entre dos pasos:

Asignación: asignamos un crimen al grupo con la media más cercana:

$$S_{(i)}^{(t)} = \left\{ x_p \,\middle|\, \|x_p - m_i^{(t)}\|_2^2 \le \|x_p - m_j^{(t)}\|_2^2 \ \forall j, 1 \le j \le k \right\}$$

cada x_p es asignado a un solo $S^{(t)}$.

Actualización: recalculamos las medias con las la media más cercana:

$$S_{(i)}^{(t)} = \{x_p \mid \|x_p - m_i^{(t)}\|_2^2 \leq \|x_p - m_j^{(t)}\|_2^2 \ \forall j, 1 \leq j \leq k\}$$

cada x_p es asignado a un solo $S^{(t)}$.

Actualización: recalculamos las medias con las observaciones asignadas a cada grupo:

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

El algoritmo converge cuando en un paso las asignaciones no cambien.