
OPTIMIZACIÓN

Primer Cuatrimestre 2015

Trabajo Práctico N° 1: Ajuste de círculos.

El objetivo de este trabajo es implementar y comparar algoritmos que ajusten conjuntos de datos con círculos, siguiendo diferentes metodologías. A lo largo de todo el texto, asumiremos que los datos vienen dados por vectores $\mathbf{x} = (x_1, \dots, x_n)^t$ e $\mathbf{y} = (y_1, \dots, y_n)^t$, de manera tal que los puntos (x_i, y_i) caigan aproximadamente en un círculo.

Ajuste algebraico

Un círculo puede definirse implícitamente con la ecuación:

$$\alpha(x^2 + y^2) + \beta x + \gamma y + \delta = 0.$$

Para garantizar unicidad en la escritura, asumiremos $\alpha = 1$. Además, según veremos más adelante, resultará útil reescribir los parámetros, notando:

$$x^2 + y^2 - 2ax - 2by - c = 0.$$

Si todos los datos cayeran sobre un círculo tendríamos que la ecuación se satisface para todo (x_i, y_i) . Asumiendo que esto no sucede (los datos contienen error), buscaremos valores de a , b y c que minimicen:

$$\sum_i (x_i^2 + y_i^2 - 2ax_i - 2by_i - c)^2.$$

O, equivalentemente:

$$\min_{a,b,c} \|2a\mathbf{x} + 2b\mathbf{y} + c - \mathbf{p}\|, \quad \text{donde } \mathbf{p} = (x_1^2 + y_1^2, \dots, x_n^2 + y_n^2)^t. \quad (1)$$

Este enfoque se conoce como *ajuste algebraico* del círculo, dado que se propone ajustar los parámetros de la ecuación algebraica que lo define.

Ejercicio 1 Observar que (1) es un problema de cuadrados mínimos y escribir la matriz del ajuste. Implementar un algoritmo que reciba como datos dos vectores \mathbf{x} e \mathbf{y} y resuelva (1) utilizando la descomposición QR implementada en el Ejercicio 18 b) de la Práctica 1.

Ajuste geométrico

El ajuste algebraico, al ceñirse a un modelo de cuadrados mínimos clásico, puede resolverse de manera simple y veloz. Sin embargo, no tiene en cuenta la naturaleza geométrica del problema: no minimiza ninguna distancia. Esto hace que en algunos casos se obtengan resultados muy

distintos de los esperados. El *ajuste geométrico* consiste en minimizar las distancias de los datos al círculo.

Notamos $C(a, b, r)$ al círculo con centro (a, b) y radio r y

$$d_i = \sqrt{(x_i - a)^2 + (y_i - b)^2} - r,$$

la distancia de (x_i, y_i) a $C(a, b, r)$. El problema de determinar los parámetros a, b, r que mejor ajusten los datos \mathbf{x}, \mathbf{y} está dado por:

$$\min F(a, b, r) \tag{2}$$

donde

$$F(a, b, r) = \sum_{i=1}^n d_i^2 = \sum_{i=1}^n \left(\sqrt{(x_i - a)^2 + (y_i - b)^2} - r \right)^2.$$

Es importante observar que este es un problema de cuadrados mínimos *no lineal*, por lo cual no es posible aplicar de manera directa las técnicas habituales de cuadrados mínimos.

En una primera aproximación podríamos modificar el funcional de manera que resulte posible aplicar las técnicas usuales de cuadrados mínimos (lineales). Esto puede lograrse, por ejemplo, considerando:

$$\tilde{F}(a, b, r) = \sum_{i=1}^n \tilde{d}_i^2 = \sum_{i=1}^n \left((x_i - a)^2 + (y_i - b)^2 - r^2 \right)^2 \tag{3}$$

Este enfoque, sin embargo, no aporta ningún avance significativo:

Ejercicio 2 Probar que minimizar \tilde{F} es equivalente a resolver el ajuste algebraico (1).

Para resolver (2) utilizaremos el método iterativo de Gauss-Newton, que describimos a continuación. El objetivo es minimizar

$$\sum_{i=1}^n d_i(\theta)^2, \quad \theta = (a, b, r).$$

Notamos $\mathbf{d}(\theta) = (d_1(\theta), \dots, d_n(\theta))^t$. Idealmente (si los datos coincidieran en un mismo círculo) tendríamos que la solución óptima θ^* satisface $\mathbf{d}(\theta^*) = \mathbf{0}$.

Desarrollando el polinomio de Taylor a orden 1 tenemos que:

$$\mathbf{d}(\theta + \mathbf{h}) \simeq \mathbf{d}(\theta) + J(\theta)\mathbf{h} \sim \mathbf{0},$$

donde J es la matriz diferencial de \mathbf{d} . El método de Gauss-Newton toma una solución inicial θ_0 , resuelve el problema de cuadrados mínimos:

$$J(\theta_0)\mathbf{h} = -\mathbf{d}(\theta_0), \tag{4}$$

toma $\theta_1 = \theta_0 + \mathbf{h}$, e itera el procedimiento, calculando una sucesión θ_i y deteniéndose cuando se satisface algún criterio de convergencia.

Ejercicio 3 Estudiar, analíticamente, la iteración que daría el método de Newton-Raphson para el problema (2). Comparar con el método de Gauss-Newton descrito anteriormente. ¿Qué ventajas presenta Gauss-Newton?

Ejercicio 4 Calcular analíticamente $J(\theta)$ para el problema (2) e implementar el algoritmo de Gauss-Newton, resolviendo (4) a través de la descomposición en valores singulares de J , $J = U\Sigma V^t$. (Usar el comando `svd` de Matlab).

Datos singulares

Cuando la matriz Σ es casi singular el error numérico del método anterior crece, haciendo que la solución resulte poco confiable. Es importante remarcar que este inconveniente *no* proviene del uso de la descomposición en valores singulares, sino que es intrínseco a la matriz J y, más esencialmente, a los datos \mathbf{x} e \mathbf{y} . El hecho de que Σ sea casi singular es un síntoma de que J está cerca de las matrices de rango menor que 3, lo que hace que el problema (4) no tenga solución única. Para resolver este inconveniente Levenberg y Marquardt propusieron una modificación del algoritmo de Gauss-Newton. La idea es la siguiente: se elige un parámetro λ y se resuelve (4) reemplazando:

$$\Sigma = \Sigma + \lambda I.$$

De este modo se *corrige* Σ para salvar su (casi) singularidad. Cuanto mayor es el valor de λ , más se parece el vector \mathbf{h} resultante al dado por el método del gradiente; cuanto menor es λ , más parecida es la iteración a la correspondiente al método de Gauss-Newton puro. La elección de λ no resulta en absoluto trivial. Marquardt sugirió una mecánica iterativa, que actualiza el valor de λ en cada paso, según el comportamiento de J :

Se toma un valor inicial λ_0 (por ejemplo $\lambda_0 = 1$) y un parámetro de corrección $\alpha > 1$. Se considera la solución inicial $\theta = (a, b, r)$ y se calculan los valores $\theta' = (a', b', r')$ y $\theta'' = (a'', b'', r'')$, correspondientes a realizar una iteración del método con $\lambda = \lambda_0$ y con $\lambda = \frac{\lambda_0}{\alpha}$, respectivamente. Para todas estas soluciones se computa el valor del funcional F . Si tanto θ' como θ'' arrojan valores de F *mayores* que el correspondiente a θ , entonces se agranda λ iterando: $\lambda = \alpha\lambda$ hasta hallar un λ que mejore el valor del funcional. Si, en cambio, $F(\theta'') < F(\theta)$, se actualiza $\lambda = \frac{\lambda}{\alpha}$, mientras que si $F(\theta') < F(\theta)$, se conserva el valor de λ actual.

Ejercicio 5 Implementar un programa que ejecute el método de Gauss-Newton en la variante de Levenberg-Marquardt.

Para probar todos los algoritmos implementados, en la página de la materia está disponible el archivo `datos.mat`. Al cargar `datos.dat` en Matlab se generan varias matrices de nombre: `datos_i` ($i = 1, \dots, 10$). Cada una de estas matrices de $n_i \times 2$ corresponde a un ejemplo y contiene en su primer columna los valores de \mathbf{x} y en la segunda los valores de \mathbf{y} .

Yapa

Ejercicio 6 Probar los algoritmos con los datos propuestos.

- ¿Qué se observa? ¿Para qué conjuntos de datos el ajuste algebraico resulta *acceptable*? ¿Qué particularidad tienen los datos cuyo ajuste *no* es bueno?
- Observar que en (1) se está minimizando $\sum_i (r_i^2 - r^2) = \sum_i (r_i - r)(r_i + r)$, donde r_i es la distancia del i -ésimo dato al centro y r es el radio del círculo. La solución al problema deberá equilibrar los aportes de los dos factores involucrados en esta suma.
- Dado un conjunto de puntos, ¿cómo conviene elegir el centro del círculo para minimizar el factor $(r_i + r)$? (No hace falta hacer cuentas: se espera una respuesta cualitativa).
- ¿Qué explicación podría darse para el ajuste algebraico en los casos en que da mal?