



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

ALGUNOS MÉTODOS PARA AJUSTAR ELIPSES

T E S I S
QUE PARA OBTENER EL TÍTULO DE:
A C T U A R I O
P R E S E N T A:
BENJAMÍN FIGUEROA SOLANO

DIRECTOR DE TESIS: DR. PABLO BARRERA SÁNCHEZ



m. 3410915



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

Autorizo a la Dirección General de Bibliotecas de la UNAM a diligenciar en formato electrónico el contenido de mi trabajo escrito.

NOMBRE: Benjamín Figueroa Solano

FECHA: 10/02/2005

FIRMA: [Firma manuscrita]

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

"Algunos métodos para ajustar elipses"

realizado por Benjamín Figueroa Solano

con número de cuenta 09620419-9 , quien cubrió los créditos de la carrera de:
 Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
 Propietario

Dr. Pablo Barrera Sánchez

Propietario

M. en A.P. María del Pilar Alonso Reyes

Propietario

M. en C. Claudia Alquicira Esquivel

Suplente

M. en C. Guilmer Ferdinand González Flores

Suplente

Mat. Luis Alberto Vázquez Maisón

Consejo Departamental de Matemáticas

Act. Jaime Vázquez Alamilla

FACULTAD DE CIENCIAS

CONSEJO DEPARTAMENTAL

MATEMÁTICAS

[Firma manuscrita]

[Firma manuscrita]

[Firma manuscrita]

[Firma manuscrita]

[Firma manuscrita]

Dedicatoria

Dedico este trabajo:

*A mis padres, Esther y Felipe,
y a mis hermanos, Armando, Graciela,
Daniel, Ma. Esther y Lourdes
con mucho cariño.*

Agradecimientos

Al Dr. Pablo Barrera Sánchez, por todo el apoyo y experiencia puestos en la dirección de este trabajo, también por su paciencia y valiosos comentarios para mejorarlo, pero sobre todo por el gran aprendizaje que me llevo de él.

A la M. en A.P. María del Pilar Alonso Reyes y a la M. en C. Claudia Alquicira Esquivel, por sus valiosos comentarios para mejorar este trabajo.

Al M. en C. Guilmer F. González Flores y al Mat. Luis Alberto Vazquéz Maison, por su apoyo, paciencia, asesoría y atinados comentarios que ayudaron a mejorar este trabajo.

A mi padre, porque me has enseñado más cosas de las que te imaginas.

A mi madre, por todos estos años que has cuidado de mí.

A mis hermanos, Armando, Graciela, Daniel, María Esther y Lourdes por su apoyo y motivación, pero sobre todo por brindarme tanto cariño.

A Eloina, por todo el amor que me has dado.

A Dios, por haberme permitido cumplir un sueño.

Contenido

Prefacio	1
1 Representaciones de la elipse	3
1.1 Introducción	3
1.2 Curvas paramétricas e implícitas	5
1.3 Diversas representaciones de la elipse	5
1.3.1 Familia de elipses	7
1.4 Representación matricial de las cónicas	8
1.5 Algunas cantidades invariantes	10
2 Ajuste de elipses	13
2.1 Introducción	13
2.2 Planteamiento del problema	13
2.3 Criterios de ajuste	17
2.4 Algunas métricas	18
2.5 Técnicas de ajuste	19
2.5.1 Ajuste por distancia algebraica	19
2.5.2 Ajuste por distancia geométrica	30
3 Resultados experimentales	37
3.1 Introducción	37
3.2 Comparación de los métodos	38
3.2.1 Invarianza	38

3.2.2	Especificidad	49
3.2.3	Excentricidad	54
3.2.4	Ruido	68
Conclusiones		83
A Mínimos cuadrados		85
A.1	Introducción	85
A.2	Ajuste lineal por mínimos cuadrados	86
A.3	Las ecuaciones normales	90
A.4	La factorización QR	94
B Mínimos cuadrados no lineales		97
B.1	Introducción	97
B.2	Solución numérica del problema	98
B.2.1	El método de Gauss-Newton	98
B.2.2	El método de Levenberg-Marquardt	99
C Sistema interactivo de ajuste de elipses		101
C.1	Introducción	101
C.2	Descripción del sistema	102
C.3	Ejemplos	107
C.3.1	Comparando los métodos <i>Paton</i> y <i>Fitzgibbon</i>	107
C.3.2	Ajuste en forma interactiva	110
C.4	Descripción de los m-files empleados	112
C.5	Distribución del código	116
Bibliografía		117

Prefacio

En esta tesis se estudia el problema del ajuste de elipses. El ajuste de elipses es una herramienta indispensable en áreas como el diseño asistido por computadora, reconocimiento de patrones y visión por computadora.

En el **primer Capítulo** comenzamos hablando sobre las cónicas y su importancia en el desarrollo de la ciencia. Posteriormente, se estudian las diversas representaciones de la elipse y se demuestra la invarianza bajo transformaciones euclidianas de algunos elementos que determinan una cónica.

En el **segundo Capítulo** planteamos el problema del ajuste de elipses y mencionamos algunas distancias a usar para la solución de este. También se revisan seis técnicas para el ajuste de elipses: cinco para el ajuste por distancia algebraica y una para el ajuste por distancia geométrica.

En el **tercer Capítulo** se llevan a cabo diversos experimentos ajustando elipses con cada uno de los métodos del segundo capítulo. Se observa el comportamiento de estos cuando los puntos que se van a ajustar son transformados o perturbados con ruido gaussiano, entre otros experimentos. También se comparan algunas características geométricas de las elipses que se obtienen al minimizar la distancia algebraica con las que se obtienen al minimizar la distancia geométrica de los puntos a la curva.

En el **Apéndice A** se trata el tema de los mínimos cuadrados y se incluyen algunos ejemplos de ajuste de polinomios con este método.

En el **Apéndice B** se abordan los mínimos cuadrados no lineales y se revisan brevemente los métodos de Gauss-Newton y Levenberg-Marquardt.

Para llevar a cabo los experimentos se desarrolló un sistema interactivo para el ajuste de elipses en MATLAB (versión 5.3.0.10183 (R11)) [22]. Un ejemplo de como usar este sistema se describe en el **Apéndice C**.

Capítulo 1

Representaciones de la elipse

1.1 Introducción

En este capítulo se revisan algunas propiedades geométricas de las cónicas, dedicándole particular atención a aquellas que tengan que ver con la elipse, ya que este objeto matemático es el motivo de la presente investigación.

El estudio de las cónicas tiene su origen en el libro *Cónicas* de Apolonio de Perga (262–192 a.C. aprox.). En su libro, Apolonio estudia de manera rigurosa las curvas que se obtienen al cortar un cono por diversos planos. Previo al trabajo de Apolonio, existían algunos resultados elementales sobre determinadas intersecciones de planos perpendiculares a las generatrices de un cono, obteniéndose elipses, parábolas e hipérbolas según el ángulo superior del cono fuese agudo, recto u obtuso, respectivamente.

Si bien faltaban algunos siglos para que se desarrollara la geometría analítica, Apolonio hace un tratamiento de las cónicas que se aproxima mucho al que se deriva de ésta. Los resultados obtenidos por él, fueron los únicos que existieron hasta que Fermat y Descartes, en una de las primeras aplicaciones de la geometría analítica, retomaron el problema llegando a estudiarlas casi por completo, salvo porque no manejaban coordenadas negativas.

Las cónicas, se obtienen al intersectar una *superficie cónica* con un plano. Se le llama *superficie cónica de revolución* a la superficie generada por una línea recta que gira alrededor de un eje, manteniendo un punto fijo sobre dicho eje; mientras que se denomina *sección cónica*¹ a la curva que se obtiene al cortar esa superficie cónica con un plano. Las diferentes posiciones de dicho

¹Un estudio completo de las secciones cónicas se puede encontrar en [5, 21, 25, 26, 30].

plano determinan distintas curvas, a saber: *circunferencia, elipse, hipérbola o parábola*.

La importancia fundamental de las cónicas radica en su constante aparición en situaciones reales:

- En el diseño asistido por computadora (CAD) permiten la representación de piezas mecánicas así como la elaboración de planos arquitectónicos o de ingeniería civil.
- Se pueden utilizar para aproximar curvas complicadas con mayor precisión, por ejemplo usando un spline cónico.
- La *Primera Ley de Kepler* sobre el movimiento de los planetas, dice que estos siguen órbitas elípticas en uno de cuyos focos se encuentra el Sol.
- Es muy posible que Newton no hubiese podido descubrir su famosa *Ley de la Gravitación Universal* de no haber conocido ampliamente la geometría de las elipses.
- La órbita que sigue un objeto dentro de un campo gravitacional constante es una parábola. Así, la línea que describe cualquier móvil que es lanzado con una cierta velocidad inicial, que no sea vertical, se puede modelar por una parábola.
- La *propiedad focal* de las cónicas es de gran utilidad en óptica y en otras aplicaciones de la física. Esta propiedad dice lo siguiente: "*Si un rayo de luz pasa por un foco, al reflejarse en la cónica lo hace siguiendo una recta que pasa por el otro foco*".

Las cónicas han sido ampliamente estudiadas y se han utilizado en una gran diversidad de aplicaciones; actualmente, se siguen usando de forma extensiva en la matemática aplicada [3, 24, 20].

Las cónicas, desde el punto de vista de una curva, pueden representarse en más de una forma; de igual manera, las elipses cuentan con diversas representaciones. Veamos algunas de ellas.

1.2 Curvas paramétricas e implícitas

Una curva se puede representar tanto en forma paramétrica como en forma implícita. Una curva parametrizada en \mathbb{R}^2 se puede ver como una colección de puntos de la forma

$$C(\mathbf{a}) = \{c(\mathbf{a}; t) \mid t_{min} \leq t \leq t_{max}\} \quad (1.1)$$

donde \mathbf{a} es un vector de parámetros que determinan a la curva; $c(\mathbf{a}; t)$ es una función vectorial

$$c(\mathbf{a}; t) = \begin{pmatrix} c_x(\mathbf{a}; t) \\ c_y(\mathbf{a}; t) \end{pmatrix} \quad (1.2)$$

y t_{min}, t_{max} son los extremos del intervalo de valores admisibles para t .

Por otra parte, una curva implícita es la colección de puntos de la forma

$$C(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^2 \mid F(\mathbf{a}; \mathbf{x}) = 0\} \quad (1.3)$$

donde $F(\mathbf{a}; \mathbf{x})$ es una función escalar en \mathbf{x} definida por el vector de parámetros \mathbf{a} . En particular, las secciones cónicas, se representan naturalmente por un polinomio cuadrático en x y y .

1.3 Diversas representaciones de la elipse

Las representaciones implícitas más conocidas de la elipse son las siguientes:

$$C(a, b) = \left\{ (x, y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad a, b \in \mathbb{R} - \{0\} \right\} \quad (1.4)$$

o bien,

$$C(a, b, h, k) =$$

$$\left\{ (x, y) \mid \frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \quad a, b \in \mathbb{R} - \{0\}, \quad h, k \in \mathbb{R} \right\} \quad (1.5)$$

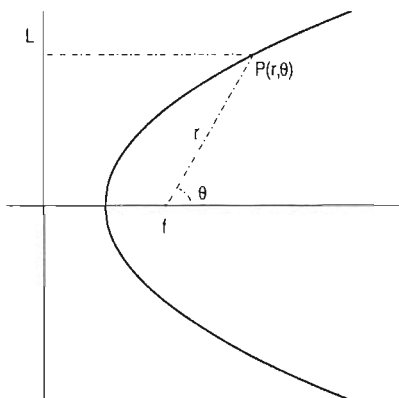
donde h y k son la abscisa y la ordenada del centro respectivamente, mientras que a y b son las longitudes de los semiejes. Presentan el inconveniente de que los parámetros a, b, h y k aparecen en forma no lineal. Estas representaciones resultan bastante familiares para el lector con nociones elementales de geometría analítica.

Otra representación muy conocida es la que se obtiene parametrizando por un ángulo θ . Dada una elipse con centro en (h, k) y semiejes a y b sus ecuaciones paramétricas son:

$$c(a, b, h, k; \theta) = \left\{ \left(\begin{array}{l} h + a \cos \theta \\ k + b \sin \theta \end{array} \right) \mid \theta \in (0, 2\pi], \quad a, b \in \mathbb{R} - \{0\} \quad h, k \in \mathbb{R} \right\} \quad (1.6)$$

Una representación de las cónicas, es la que resulta de considerar una recta fija L y un punto fijo f fuera de la recta; se quiere estudiar el lugar geométrico C_e de los puntos P del plano tales que su distancia al punto fijo f , llamado *foco*, y a la recta fija L , llamada *directriz*, tiene una razón constante e (necesariamente positiva), la cual es la *excentricidad* de la cónica. Es decir,

$$C_e = \{P \mid d(P, f) = ed(P, L)\} \quad (1.7)$$



De esta manera, se puede caracterizar a la elipse en términos de la excentricidad, tal como lo enuncia el siguiente criterio.

Criterio de la excentricidad. *Una cónica es elipse si su excentricidad es menor que 1.*

La forma polar de las cónicas se obtiene al poner el polo en uno de los focos, la directriz a una distancia p del foco y (r, θ) las coordenadas polares de P , con lo cual:

$$C(e, p) = \left\{ (r, \theta) \mid r = \frac{ep}{1 - e \cos \theta} \quad e, p \in \mathbb{R}^+, \quad \theta \in [0, \pi] \right\} \quad (1.8)$$

Las coordenadas polares facilitan ciertos cálculos cuando se utilizan en el contexto adecuado.

1.3.1 Familia de elipses

La familia de cónicas

$$C(A, B, C, D, E, F) =$$

$$\{(x, y) \mid Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0\} \quad (1.9)$$

es de las más usadas en la práctica debido a que los parámetros A, B, C, D, E y F aparecen en forma lineal. El criterio usual para saber cuando la cónica es una elipse es el siguiente:

Criterio del discriminante. *En la ecuación general de segundo grado $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$, si el discriminante $B^2 - 4AC$ es negativo ésta representa una elipse.*

Si A y C son positivos es posible visualizar geoméricamente cuándo la cónica (1.9) representa una elipse. Si se suponen A y C positivos y se hace $K = A+C$ entonces $K - A = C > 0$, substituyendo en el discriminante resulta

$$B^2 - 4A(K - A)$$

entonces, de la desigualdad

$$B^2 - 4A(K - A) \leq 0$$

haciendo un poco de álgebra se llega a

$$\frac{B^2}{K^2} + \frac{(A - K/2)^2}{(K/2)^2} \leq 1 \quad (1.10)$$

que corresponde a una elipse con centro en $(K/2, 0)$, semieje mayor K y semieje menor $K/2$, más todo el interior. Observe la figura (1.1).

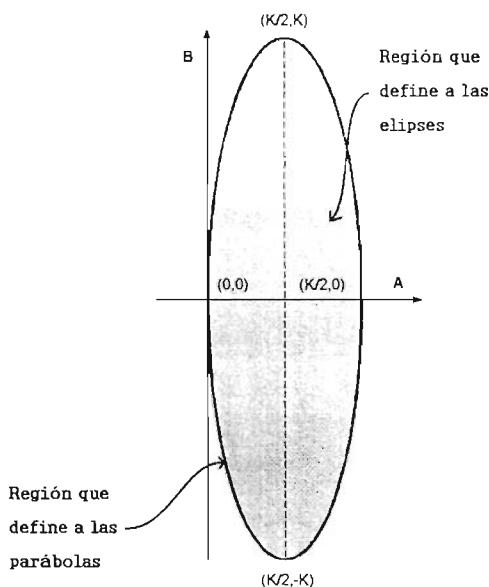


Figura 1.1: Región en (1.10)

Si A y C son negativos y se sigue el procedimiento anterior se llega a la misma desigualdad. Entonces se tiene una elipse cuando (A, B) es un punto del interior de (1.10), y si (A, B) está en la frontera se obtiene una parábola.

Una representación más es la matricial, la cual será tratada en la siguiente sección.

1.4 Representación matricial de las cónicas

La ecuación general de segundo grado

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

se puede escribir de la forma

$$\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{x}^T \mathbf{b} + F = 0$$

donde

$$\mathbf{Q} = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} D \\ E \end{pmatrix}$$

con esto la cónica queda representada de la siguiente manera

$$C(\mathbf{Q}, \mathbf{b}, F) = \{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{x}^T \mathbf{b} + F = 0 \} \quad (1.11)$$

Considere una transformación euclídeana de la forma $\mathbf{x}' = M\mathbf{x} + \mathbf{t}$, donde M es una rotación y \mathbf{t} una traslación. Con esto M satisface que $M^T M = M M^T = \mathbf{I}_{2 \times 2}$, entonces esta matriz cumple con ser ortogonal² y por consiguiente $M^{-1} = M^T$.

Aplicando esta transformación a la cónica se obtiene

$$\mathbf{x}'^T \mathbf{Q}' \mathbf{x}' + \mathbf{x}'^T \mathbf{b}' + F' = 0 \quad (1.12)$$

Para relacionar los parámetros transformados con los originales, se expande en términos de los puntos originales \mathbf{x}

$$(M\mathbf{x} + \mathbf{t})^T \mathbf{Q}' (M\mathbf{x} + \mathbf{t}) + (M\mathbf{x} + \mathbf{t})^T \mathbf{b}' + F' = 0$$

Reuniendo los términos en \mathbf{x} y comparando con la ecuación (1.12), se tiene que

$$\begin{aligned} \mathbf{Q} &= M^T \mathbf{Q}' M \\ \mathbf{b} &= M^T (2\mathbf{Q}' \mathbf{t} + \mathbf{b}') \\ F &= \mathbf{t}^T \mathbf{Q}' \mathbf{t} + \mathbf{t}^T \mathbf{b}' + F' \end{aligned}$$

Resolviendo estas ecuaciones se obtienen expresiones para los parámetros transformados en términos de los originales:

$$\mathbf{Q}' = M^{-T} \mathbf{Q} M^{-1} \quad (1.13)$$

$$\mathbf{b}' = M^{-T} \mathbf{b} - 2\mathbf{Q}' \mathbf{t} \quad (1.14)$$

$$F' = \mathbf{t}^T \mathbf{Q}' \mathbf{t} - \mathbf{t}^T M^{-T} \mathbf{b}' + F \quad (1.15)$$

Obsérvese que bajo esta representación el discriminante se puede escribir como un múltiplo del determinante de la matriz \mathbf{Q}

$$-4 \det(\mathbf{Q}) = B^2 - 4AC \quad (1.16)$$

Las cantidades

$$\text{traza}(\mathbf{Q}) = A + C \quad \text{y} \quad \text{traza}(\mathbf{Q}^2) = A^2 + B^2/2 + C^2$$

serán estudiadas más adelante y se verá que, al igual que el discriminante, éstas también son invariantes bajo transformaciones euclídeanas.

²Una matriz cuadrada \mathbf{A} es ortogonal si $\mathbf{A} \mathbf{A}^T = \mathbf{I}$

1.5 Algunas cantidades invariantes

En la sección anterior se vio que

$$\mathbf{Q} = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \quad y \quad \mathbf{Q}' = M^{-T} \mathbf{Q} M^{-1}$$

donde la matriz \mathbf{Q}' contiene a los coeficientes A', B', C' que resultan de la rotación, pero además

$$\mathbf{Q}' = M^{-T} \mathbf{Q} M^{-1} = M \mathbf{Q} M^T \quad (1.17)$$

Como M es ortogonal el $\det(M) = \pm 1$, de donde

$$\det(\mathbf{Q}') = \det(M \mathbf{Q} M^T) = \det^2(M) \det(\mathbf{Q}) = \det(\mathbf{Q})$$

con lo que la invarianza del discriminante queda probada.

De la siguiente igualdad

$$\mathbf{Q}' - \lambda \mathbf{I} = M \mathbf{Q} M^T - \lambda M M^T = M(\mathbf{Q} - \lambda \mathbf{I}) M^T$$

es claro que

$$\det(\mathbf{Q}' - \lambda \mathbf{I}) = \det(\mathbf{Q} - \lambda \mathbf{I})$$

y por lo tanto \mathbf{Q} y \mathbf{Q}' tienen los mismos eigenvalores. Como la suma de los eigenvalores de una matriz es igual a la traza de la misma se sigue que $\text{traza}(\mathbf{Q}) = A + C$ es invariante.

De igual forma para \mathbf{Q}'^2 y \mathbf{Q}^2 se tiene que

$$\begin{aligned} (\mathbf{Q}')^2 - \lambda \mathbf{I} &= (M \mathbf{Q} M^T)(M \mathbf{Q} M^T) - \lambda \mathbf{I} \\ &= M \mathbf{Q}^2 M^T - \lambda M M^T \\ &= M(\mathbf{Q}^2 - \lambda \mathbf{I}) M^T \end{aligned}$$

y entonces si se usan los mismos argumentos de antes \mathbf{Q}'^2 y \mathbf{Q}^2 tienen los mismos eigenvalores y por lo tanto $\text{traza}(\mathbf{Q}'^2) = \text{traza}(\mathbf{Q}^2)$, i.e. $A^2 + B^2/2 + C^2$ es invariante.

A diferencia de lo que se vio para $A + C$, $B^2 - 4AC$ y $A^2 + B^2/2 + C^2$ algunas cantidades no son invariantes bajo transformaciones euclidianas.

Considere la elipse

$$2x^2 + xy + 2y^2 + 3x + 3y - 1 = 0 \quad (1.18)$$

Si se aplica una rotación de $\theta = \pi/6$ y una traslación por $(3, -3)$ de las igualdades (1.13), (1.14) y (1.15) se obtienen los parámetros transformados

$$\begin{aligned} Q' &= \frac{1}{8} \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix} \\ &= \begin{pmatrix} 2 - \sqrt{3}/4 & 1/4 \\ 1/4 & 2 + \sqrt{3}/4 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{b}' &= \frac{1}{2} \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \begin{pmatrix} 3 \\ 3 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 8 - \sqrt{3} & 1 \\ 1 & 8 + \sqrt{3} \end{pmatrix} \begin{pmatrix} 3 \\ -3 \end{pmatrix} \\ &= \begin{pmatrix} 3\sqrt{3} - 12 \\ 3\sqrt{3} + 12 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} F' &= \begin{pmatrix} 3 & -3 \end{pmatrix} \begin{pmatrix} 2 - \sqrt{3}/4 & 1/4 \\ 1/4 & 2 + \sqrt{3}/4 \end{pmatrix} \begin{pmatrix} 3 \\ -3 \end{pmatrix} \\ &\quad - \begin{pmatrix} 3 & -3 \end{pmatrix} \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} \begin{pmatrix} 3\sqrt{3} - 12 \\ 3\sqrt{3} + 12 \end{pmatrix} - 1 \\ &= 61/2 + 45\sqrt{3} \end{aligned}$$

Con los coeficientes de (1.18) se evalúa la cantidad

$$A^2 + B^2 + C^2 + D^2 + E^2 + F^2 = 2^2 + 1^2 + 2^2 + 3^2 + 3^2 + (-1)^2 = 28$$

por otra parte, con los coeficientes transformados esta cantidad vale

$$\begin{aligned} A'^2 + B'^2 + C'^2 + D'^2 + E'^2 + F'^2 &= (2 - \sqrt{3}/4)^2 + (1/2)^2 + (2 + \sqrt{3}/4)^2 \\ &\quad + (3\sqrt{3} - 12)^2 + (3\sqrt{3} + 12)^2 \\ &\quad + (61/2 + 45\sqrt{3})^2 \\ &= 12110.35 \end{aligned}$$

entonces, $\| (A, B, C, D, E, F)^T \|^2$ no es invariante.

Otra cantidad que también es no invariante es F . En la ecuación (1.18) se tiene que $F = -1$ mientras que bajo las transformaciones $F' = 61/2 + 45\sqrt{3}$.

Como se ha visto hasta ahora son muy pocas las cantidades que si son invariantes. De hecho, casi ninguna función de los parámetros de la ecuación general de segundo grado será invariante bajo transformaciones euclidianas, sin embargo, las que si lo son, jugarán un papel importante en el ajuste de elipses.

Capítulo 2

Ajuste de elipses

2.1 Introducción

En este capítulo se estudia el ajuste de elipses por diferentes técnicas. Se revisan algunos de los métodos existentes para ajustar cónicas, que bajo ciertas condiciones, resultan apropiados para ajustar elipses.

2.2 Planteamiento del problema

Cinco puntos en el plano determinan una cónica. Así pues, si se tienen al menos cinco puntos en el plano se puede determinar de manera única una elipse, una hipérbola o una parábola. Sin embargo, en muchas de las aplicaciones la cantidad de puntos con los que se debe trabajar resulta ser considerablemente mayor a cinco. Entonces, al contar con más de cinco puntos el problema es determinar una elipse que minimice la distancia de los puntos a la curva. La métrica natural de \mathbb{R}^2 es la distancia euclídeana, sin embargo, como se verá más adelante, existen otras formas de definir distancia en el plano cartesiano.

Este trabajo se aboca al problema de ajustar una elipse a un conjunto de puntos. Este problema se puede plantear de la siguiente manera: dado un conjunto de puntos en el plano $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, con al menos cinco elementos, se quiere determinar la elipse que mejor se ajusta a ellos. Usualmente, se elige una representación de la elipse que depende de un vector de parámetros \mathbf{a} y entonces el problema es determinar tales parámetros usando algún criterio de ajuste.

Veamos algunos ejemplos que ilustren hacia donde vamos. En la fig.

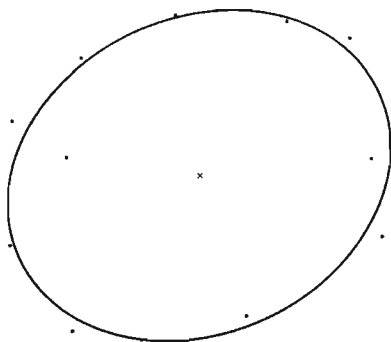


Figura 2.1: Elipse ajustando un conjunto de puntos

(2.1) se observa una elipse obtenida al ajustarla a un conjunto de puntos en el plano. Determinar los parámetros de una representación de la elipse es el problema que se abordará en este capítulo.

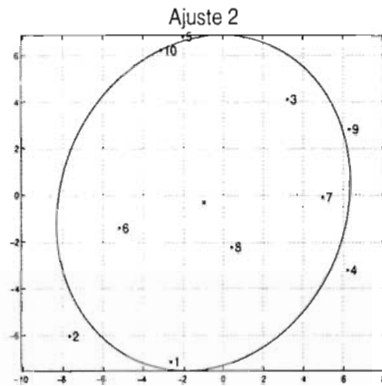
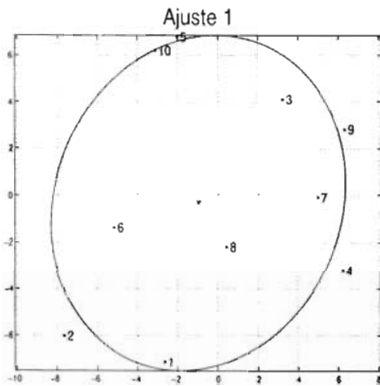
En las siguientes gráficas (Figura 2.2) se observan seis elipses que ajustan a un mismo conjunto de puntos; cada una de estas corresponde a la solución que se encuentra de aplicar un criterio de ajuste para una representación.

Para el ajuste por distancia algebraica se recurre a la ecuación general de segundo grado para representar a la elipse, mientras que en el ajuste por distancia geométrica se usa la forma paramétrica. Entonces, como se verá más adelante, la solución no depende sólo de los puntos, sino también del criterio de ajuste y la representación de la elipse que se esté usando.

Los puntos que se usaron en este ejemplo son los siguientes:

x	y
-2.672811	-7.134503
-7.695853	-6.023392
3.225806	4.093567
6.221198	-3.216374
-2.027650	6.783626
-5.207373	-1.403509
4.976959	-0.116959
0.414747	-2.222222
6.313364	2.807018
-3.133641	6.198830

Tabla 2.1: Puntos



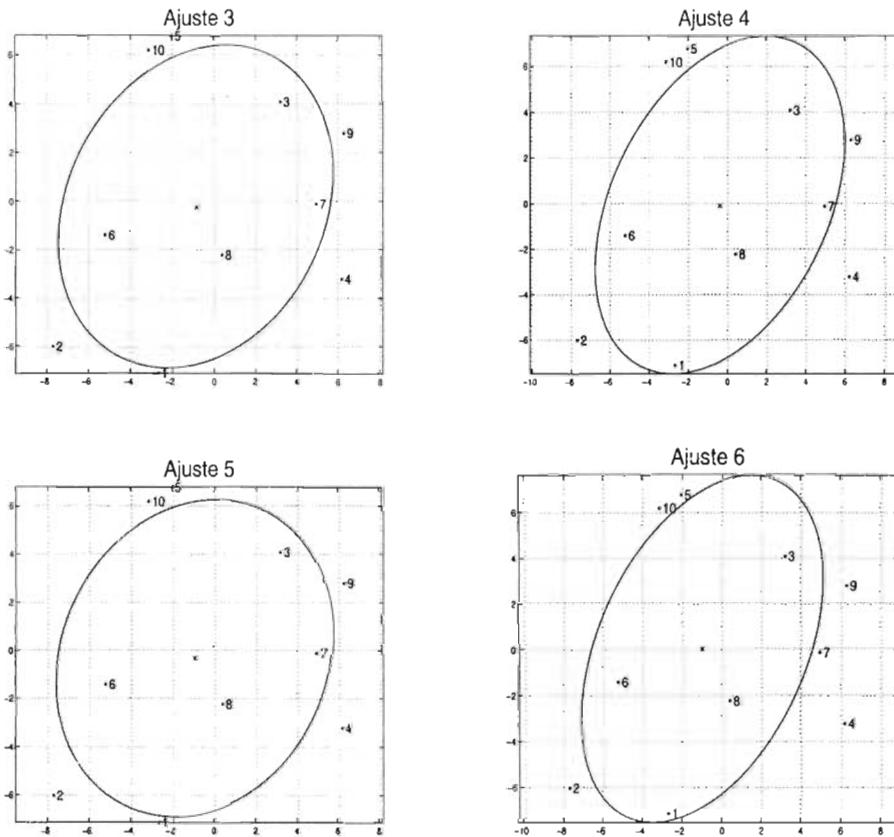


Figura 2.2: Elipses que se obtienen al ajustar un mismo conjunto de puntos usando distintos criterios y representaciones.

2.3 Criterios de ajuste

Es necesario contar con una forma de medir que tan cerca está un punto de una curva, en la práctica, es usual evitar la distancia geométrica, pues resulta ser costosa en términos computacionales. Así, lo que se hace es usar una representación $C(\mathbf{a})$ de una curva en términos de un vector de parámetros \mathbf{a} , de tal forma que

$$C(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{Q}(\mathbf{x}; \mathbf{a}) = 0\}$$

donde la representación $\mathbf{Q}(\mathbf{x}; \mathbf{a})$ dice que un punto $\mathbf{x} \notin C(\mathbf{a})$ si $\mathbf{Q}(\mathbf{x}; \mathbf{a}) \neq 0$; con lo cual, la distancia del punto \mathbf{x} a la curva $C(\mathbf{a})$ se puede definir en términos de la representación que se esté utilizando como

$$\mathcal{D}(\mathbf{x}, C(\mathbf{a})) = \varphi(\mathbf{Q}(\mathbf{x}; \mathbf{a})).$$

donde φ es un función positiva, por ejemplo $\varphi(t) = t^2$, $\varphi(t) = |t|$. Entonces, se define la distancia de la curva $C(\mathbf{a})$ al conjunto de puntos $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, como

$$\mathcal{D}(\mathbf{X}, C(\mathbf{a})) = \sum_{i=1}^n \mathcal{D}(\mathbf{x}_i; C(\mathbf{a}))$$

con lo cual, elegir la mejor curva que ajusta los datos equivale a elegir el vector óptimo \mathbf{a}^* que minimiza

$$\min_{\mathbf{a} \in \mathcal{S}_{\mathbf{a}}} \mathcal{D}(\mathbf{X}, C(\mathbf{a})) = \mathcal{D}(\mathbf{X}, C(\mathbf{a}^*))$$

donde $\mathcal{S}_{\mathbf{a}}$ es la región de valores admisibles de los parámetros.

Por ejemplo: Si

$$\mathbf{Q}(\mathbf{x}; \mathbf{a}) = \begin{bmatrix} x^2 & xy & y^2 & x & y & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \\ E \\ F \end{bmatrix} \quad (2.1)$$

algunos espacios de parámetros comúnmente usados para (2.1) son:

a) $\mathcal{S}_{\mathbf{a}} = \{\mathbf{a} \in \mathbb{R}^6 \mid \mathbf{a} \neq 0\}$

b) $\mathcal{S}_{\mathbf{a}} = \{\mathbf{a} \in \mathbb{R}^6 \mid \mathbf{a}^T \mathbf{e}_6 = 1, \quad \mathbf{e}_6 = (0, 0, 0, 0, 0, 1) \neq 0\}$

c) $\mathcal{S}_a = \{\mathbf{a} \in \mathbb{R}^6 \mid \mathbf{a}^T(\mathbf{e}_1 + \mathbf{e}_3) = 1\}$

d) $\mathcal{S}_a = \{\mathbf{a} \in \mathbb{R}^6 \mid \|\mathbf{a}\|^2 = 1\}$

e) $\mathcal{S}_a = \left\{ \mathbf{a} \in \mathbb{R}^6 \mid \mathbf{a}^T S_B \mathbf{a} = 1, \right.$

$$S_B = \begin{pmatrix} S_b & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{pmatrix}, \quad S_b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left. \right\}$$

f) $\mathcal{S}_a = \left\{ \mathbf{a} \in \mathbb{R}^6 \mid \mathbf{a}^T S_F \mathbf{a} = 1, \right.$

$$S_F = \begin{pmatrix} S_f & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{pmatrix}, \quad S_f = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{pmatrix} \left. \right\}$$

2.4 Algunas métricas

Algunas métricas comúnmente usadas son las siguientes:

- La *distancia algebraica*. Las curvas implícitas introducen una distancia métrica natural desde cualquier punto sobre el plano a la curva. La distancia algebraica del punto \mathbf{x} a la curva definida por $\mathbf{Q}(\mathbf{x}; \mathbf{a}) = 0$ es simplemente el valor de $\mathbf{Q}(\mathbf{x}; \mathbf{a})$ en \mathbf{x}

$$\mathcal{D}_A(\mathbf{x}, C(\mathbf{a})) = |\mathbf{Q}(\mathbf{x}; \mathbf{a})|^2 \tag{2.2}$$

La principal ventaja de esta distancia métrica es que simplifica enormemente los cálculos en muchos algoritmos y por esta razón es ampliamente usada.

- La *distancia geométrica*. Es la distancia entre el punto y su punto más cercano sobre la curva

$$\mathcal{D}_G(\mathbf{x}, C(\mathbf{a})) = \min_{\mathbf{p} \in C(\mathbf{a})} \|\mathbf{p} - \mathbf{x}\| \tag{2.3}$$

Esta métrica, es frecuentemente citada por los autores como la medida óptima a minimizar desde un punto de vista de mínimos cuadrados, pero por lo general es computacionalmente costosa.

- La *distancia gradiente*. También conocida como distancia de Sampson [28], se define para curvas implícitas $Q(\mathbf{x}; \mathbf{a}) = 0$ como

$$D_{\nabla}(\mathbf{x}, C(\mathbf{a})) = \frac{|Q(\mathbf{x}; \mathbf{a})|}{\|\nabla Q(\mathbf{x}; \mathbf{a})\|} \quad (2.4)$$

Se deriva como una aproximación de primer orden para la distancia geométrica y a pesar de que generalmente es mucho más simple de calcular, es moderadamente costosa para usarse directamente en problemas de aproximación [10]. No será utilizada en este trabajo.

2.5 Técnicas de ajuste

En esta sección, se van a ver algunas técnicas para el ajuste de elipses utilizando las métricas que se mencionaron en la sección anterior. Las métricas que se usarán en este trabajo son la distancia algebraica y la distancia geométrica.

2.5.1 Ajuste por distancia algebraica

Sean $\mathbf{a} = [A, B, C, D, E, F]$ y $\mathbf{x} = (x, y)$, entonces

$$Q(\mathbf{a}; \mathbf{x}) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F \quad (2.5)$$

es una familia de cónicas parametrizada por \mathbf{a} . De hecho la expresión en (2.5) es la misma que (1.9) que ya se estudió en la sección (1.3).

Si $\mathbf{x}_i = (x_i, y_i)$ es un punto en el plano, el valor de $Q(\mathbf{a}; \mathbf{x}_i)$ será la distancia algebraica del punto \mathbf{x}_i a la cónica determinada por el vector de parámetros \mathbf{a} . En gran parte de la literatura sobre ajuste de cónicas se utiliza (2.5); esta representación será muy útil ya que los parámetros \mathbf{a} a ajustar se encuentran en forma lineal.

Si $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ es un conjunto de puntos en el plano, entonces la distancia algebraica de \mathbf{X} a la cónica $Q(\mathbf{x}; \mathbf{a})$ está dada por

$$D_A(\mathbf{X}, C(\mathbf{a})) = \sum_{i=1}^n Q(\mathbf{x}_i; \mathbf{a})^2 = \|\mathbf{D}_X \mathbf{a}\|^2$$

donde

$$D_X = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & 1 \end{pmatrix} \quad (2.6)$$

Así se obtiene el problema de determinar \mathbf{a}^* tal que

$$\min_{\mathbf{a} \in \mathcal{S}_a} \mathcal{D}_A(\mathbf{X}, C(\mathbf{a})) = \mathcal{D}_A(\mathbf{x}, C(\mathbf{a}^*))$$

la elección de $\mathcal{D}_A(\mathbf{x}, C(\mathbf{a}))$ como una suma de cuadrados usualmente garantiza una solución única \mathbf{a}^* . Se verá que los resultados que se obtengan van a diferir por la representación de la cónica, es decir, por la elección del espacio de parámetros \mathcal{S}_a que a veces en la literatura se le llama normalización. Puesto que $\mathbf{Q}(\mathbf{x}; \mathbf{a})$ y $k\mathbf{Q}(\mathbf{x}; \mathbf{a})$ representan la misma cónica la normalización introduce de manera natural un escalamiento de los parámetros.

- La mayoría de los autores mencionan la normalización $\|\mathbf{a}\|^2 = 1$. Al parecer el primero en proponerla fue Paton [24]. Esta restricción tiene la ventaja de que todas las cónicas pueden ser escaladas para satisfacerla, sin embargo, como se ha visto no es invariante ni siquiera para transformaciones euclidianas.
- Gnanadesikan [15] usa $A^2 + B^2 + C^2 + D^2 + E^2 = 1$, la cual excluye únicamente el caso degenerado $1 = 0$, pero de nuevo es no invariante para transformaciones euclidianas. No será tratada en este trabajo.
- La normalización $F = 1$ no permite ajustar elipses que pasen a través del origen y no es invariante bajo transformaciones rígidas; sin embargo, en [27] reportan –basados en resultados experimentales– que las variaciones bajo rotaciones son menores. Las primeras referencias sobre esta normalización de las que se tiene conocimiento son [1, 3, 6].
- La restricción $A + C = 1$ excluye únicamente a la familia de hipérbolas equiláteras $x^2 - y^2 = k$. En [27] reportan que esta normalización produce elipses más excéntricas que la normalización $F = 1$. En cuanto a sus propiedades geométricas resulta ser invariante bajo transformaciones de similitud (ya se probó para transformaciones euclidianas en la sección (1.5)).
- Bookstein [4] propone $A^2 + \frac{1}{2}B^2 + C^2 = 1$, invariante bajo transformaciones de similitud de los ejes y permite representar a todas las cónicas

pero excluye a las rectas, para las cuales $A = B = C = 0$. En la sección (1.5) se probó que esta cantidad es invariante bajo transformaciones euclidianas.

- Fitzgibbon [9] propone la restricción $4AC - B^2 = 1$, que incorpora la especificidad de elipse y resulta ser invariante bajo transformaciones euclidianas (sección (1.5)). Una de las ventajas de usar esta restricción es su alta eficiencia computacional.
- Forsyth *et al.* [11] proponen la restricción cúbica $\det(\mathbf{G}) = 1$, donde

$$\mathbf{G} = \begin{pmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{pmatrix}$$

aunque ellos afirman que la minimización sujeta a esta restricción resulta en un ajuste proyectivamente invariante, en [19] demuestran que el ajuste es de hecho únicamente afín invariante. No será considerada en este trabajo.

- La aproximación del cuadrado de la distancia de Taubin [29], también puede ser vista como la restricción cuadrática $\|N\mathbf{a}\|^2 = 1$ donde N es la matriz Jacobiana $[\nabla F(\mathbf{a}; \mathbf{x}_1) \mid \nabla F(\mathbf{a}; \mathbf{x}_2) \mid \dots \mid \nabla F(\mathbf{a}; \mathbf{x}_N)]$. No será tratada en este trabajo.

Cabe hacer la siguiente aclaración. Como ya se había mencionado la forma cuadrática $\mathbf{Q}(\mathbf{a}; \mathbf{x})$ representa una familia de cónicas parametrizada por \mathbf{a} . Lo que se hace al imponer restricciones a dicho vector es limitar la familia de cónicas a una subfamilia más pequeña. Entonces, depende de la restricción que se use es la subfamilia de la que se obtienen las cónicas. En algunos casos se obtendrán elipses pero puede ser también que se generen otro tipo de cónicas.

Así pues, la mayoría de los métodos algebraicos que se trabajan en esta tesis incluyen como un caso particular el de la elipse y es para ese caso específico que se estudia su comportamiento. El método que propone Fitzgibbon [9] resulta ser el único que presenta la especificidad de elipse. De aquí, se deriva una diferencia importante entre este último y los demás métodos algebraicos. Más adelante, cuando se comparen los métodos, se discutirá de que manera influye esta característica en el ajuste de elipses.

Normalización con $\| \mathbf{a} \|^2 = 1$

Esta normalización, también conocida como lineal [24], simplemente evita que los coeficientes de la ecuación general de segundo grado sean todos cero; en consecuencia, la solución que se obtenga puede corresponder tanto a una elipse como a una hipérbola o una parábola.

Considere el problema de

$$\min \sum_{i=1}^n Q(\mathbf{a}; \mathbf{x}_i)^2 = \min \| \mathbf{D}_X \mathbf{a} \|^2$$

sujeto a $\| \mathbf{a} \|^2 = 1$. Este problema de minimización con restricciones se puede pasar a otro sin restricciones usando los multiplicadores de Lagrange. Sea E la función

$$E = \| \mathbf{D}_X \mathbf{a} \|^2 - \lambda (\| \mathbf{a} \|^2 - 1) = \mathbf{a}^T \mathbf{D}_X^T \mathbf{D}_X \mathbf{a} - \lambda (\mathbf{a}^T \mathbf{a} - 1) \quad (2.7)$$

calculando el gradiente con respecto de \mathbf{a} de la función E se llega al problema de eigenvalores

$$\nabla_{\mathbf{a}} E = 0 \iff 2\mathbf{D}_X^T \mathbf{D}_X \mathbf{a} - 2\lambda \mathbf{a} = 0 \quad (2.8)$$

es decir,

$$\mathbf{D}_X^T \mathbf{D}_X \mathbf{a} = \lambda \mathbf{a} \quad (2.9)$$

entonces λ debe ser un eigenvalor de $\mathbf{D}_X^T \mathbf{D}_X$.

Ahora, observemos que si $\mathbf{D}_X^T \mathbf{D}_X \mathbf{a} = \lambda \mathbf{a}$ se cumple que

$$\| \mathbf{D}_X \mathbf{a} \|^2 = \mathbf{a}^T \mathbf{D}_X^T \mathbf{D}_X \mathbf{a} = \lambda \| \mathbf{a} \|^2 = \lambda$$

es decir, el valor mínimo se alcanza cuando \mathbf{a} es el eigenvector de $\mathbf{D}_X^T \mathbf{D}_X$ correspondiente al más pequeño eigenvalor.

Normalización con $F = 1$

Otra normalización comúnmente usada es $F = 1$. Rosin [27] estudia esta normalización y la compara con la normalización $A + C = 1$. Puesto que presenta singularidades para las cónicas que pasan por el origen y no es invariante bajo transformaciones euclidianas, esta normalización podría no ser tan apropiada para el ajuste de elipses. En [27], Rosin señala que el problema de la singularidad puede ser evitado si se hace una traslación de tal forma que los datos queden centrados en el origen; también menciona

que se obtienen elipses con menor excentricidad por considerar la restricción $F = 1$ que por considerar la restricción $A + C = 1$.

Si $F = 1$, el vector de parámetros \mathbf{a} es de la forma

$$\mathbf{a} = (A, B, C, D, E, 1)^T$$

y con esto el problema de minimizar

$$\min \sum_{i=1}^n \mathbf{Q}(\mathbf{a}; \mathbf{x}_i)^2 = \min \|\mathbf{D}_X \mathbf{a}\|^2$$

se puede resolver de la siguiente manera. Sea

$$\begin{aligned} \|\mathbf{D}_X \mathbf{a}\|^2 &= \left\| \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \\ 1 \end{pmatrix} \right\|^2 \\ &= \left\| \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \right\|^2. \end{aligned}$$

si se nombran $\bar{\mathbf{a}} = (A, B, C, D, E)^T$, $\bar{\mathbf{1}} = (1, \dots, 1)^T$ y $\bar{\mathbf{D}}_X$ a la matriz

$$\bar{\mathbf{D}}_X = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n \end{pmatrix}$$

entonces

$$\min \|\mathbf{D}_X \mathbf{a}\|^2 = \min \|\bar{\mathbf{D}}_X \bar{\mathbf{a}} + \bar{\mathbf{1}}\|^2$$

Se puede encontrar el $\bar{\mathbf{a}}$ que minimiza $\|\bar{\mathbf{D}}_X \bar{\mathbf{a}} + \bar{\mathbf{1}}\|^2$ resolviendo por factorización QR.

Normalización con $A + C = 1$

Esta normalización no es específica para elipses ya que la hipérbola $3x^2 - 2y^2 = 0$ satisface la restricción $A + C = 1$. De hecho, para que se obtenga una elipse de la ecuación general de segundo grado A y C deben tener el mismo signo, lo cual claramente no está contemplado en esta normalización [9]. Esta restricción ha sido estudiada por Rosin [27] y Gander [14]. Debido al ligero sesgo que presenta respecto a la elipse y a que es invariante bajo transformaciones rígidas es una de las más utilizadas. En la sección (1.5) se mostró que la cantidad $A + C$ es invariante bajo rotaciones.

Haciendo $A = 1 - C$ el vector \mathbf{a} queda de la forma

$$\mathbf{a} = (1 - C, B, C, D, E, F)^T$$

Para minimizar

$$\min \sum_{i=1}^n Q(\mathbf{a}; \mathbf{x}_i)^2 = \min \|\mathbf{D}_X \mathbf{a}\|^2$$

se procede de manera análoga al planteamiento anterior, con lo cual

$$\|\mathbf{D}_X \mathbf{a}\|^2 = \left\| \begin{pmatrix} x_1 y_1 & y_1^2 - x_1^2 & x_1 & y_1 & 1 \\ x_2 y_2 & y_2^2 - x_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n y_n & y_n^2 - x_n^2 & x_n & y_n & 1 \end{pmatrix} \begin{pmatrix} B \\ C \\ D \\ E \\ F \end{pmatrix} + \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix} \right\|^2$$

llamando $\hat{\mathbf{a}} = (B, C, D, E, F)^T$, $\hat{\mathbf{b}} = (x_1^2, \dots, x_n^2)^T$ y $\hat{\mathbf{D}}_X$ a la matriz

$$\hat{\mathbf{D}}_X = \begin{pmatrix} x_1 y_1 & y_1^2 - x_1^2 & x_1 & y_1 & 1 \\ x_2 y_2 & y_2^2 - x_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n y_n & y_n^2 - x_n^2 & x_n & y_n & 1 \end{pmatrix}$$

entonces

$$\min \|\mathbf{D}_X \mathbf{a}\|^2 = \min \|\hat{\mathbf{D}}_X \hat{\mathbf{a}} + \hat{\mathbf{b}}\|^2$$

Se puede encontrar el $\hat{\mathbf{a}}$ que minimiza $\|\hat{\mathbf{D}}_X \hat{\mathbf{a}} + \hat{\mathbf{b}}\|^2$ resolviendo por factorización QR.

Normalización con $A^2 + \frac{1}{2}B^2 + C^2 = 1$

Esta normalización fue propuesta por Bookstein [4]. También ha sido estudiada por Gander [14]. Veamos una manera novedosa de resolver el problema

$$\min \sum_{i=1}^n Q(\mathbf{a}; \mathbf{x}_i)^2 = \min \| \mathbf{D}_X \mathbf{a} \|^2$$

sujeto a $A^2 + \frac{1}{2}B^2 + C^2 = 1$ usando la descomposición de matrices por bloques.

Sea \mathbf{D}_X como en (2.6) y \mathbf{S}_B la matriz por bloques

$$\mathbf{S}_B = \begin{pmatrix} \mathbf{S}_b & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{pmatrix} \quad \text{donde} \quad \mathbf{S}_b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

entonces el problema de ajustar la cónica $Q(\mathbf{a}; \mathbf{x})$ sujeto a $A^2 + B^2/2 + C^2 = 1$ se plantea de la siguiente forma

$$\begin{aligned} \min \quad & \mathbf{a}^T \mathbf{D}_X^T \mathbf{D}_X \mathbf{a} \\ \text{s.a.} \quad & \mathbf{a}^T \mathbf{S}_B \mathbf{a} = 1 \end{aligned}$$

al descomponer al vector \mathbf{a} en sus coeficientes cuadráticos y lineales

$$\mathbf{a} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \quad \text{con} \quad \mathbf{u} = (A, B, C)^T \quad \text{y} \quad \mathbf{v} = (D, E, F)^T$$

y separar a la matriz $\tilde{\mathbf{D}} = \mathbf{D}_X^T \mathbf{D}_X$ en las submatrices

$$\tilde{\mathbf{D}} = \begin{pmatrix} \tilde{\mathbf{D}}_{11} & \tilde{\mathbf{D}}_{12} \\ \tilde{\mathbf{D}}_{12}^T & \tilde{\mathbf{D}}_{22} \end{pmatrix}$$

el problema puede ser reescrito como

$$\min_{\mathbf{u}, \mathbf{v}} \quad \mathbf{u}^T \tilde{\mathbf{D}}_{11} \mathbf{u} + 2\mathbf{u}^T \tilde{\mathbf{D}}_{12} \mathbf{v} + \mathbf{v}^T \tilde{\mathbf{D}}_{22} \mathbf{v} \quad (2.10)$$

$$\text{s.a.} \quad \mathbf{u}^T \mathbf{S}_b \mathbf{u} = 1 \quad (2.11)$$

ahora, la función a minimizar por multiplicadores de Lagrange es la siguiente

$$E(\mathbf{u}, \mathbf{v}, \lambda) = \mathbf{u}^T \tilde{\mathbf{D}}_{11} \mathbf{u} + 2\mathbf{u}^T \tilde{\mathbf{D}}_{12} \mathbf{v} + \mathbf{v}^T \tilde{\mathbf{D}}_{22} \mathbf{v} - \lambda(\mathbf{u}^T \mathbf{S}_b \mathbf{u} - 1)$$

calculando las derivadas parciales de E con respecto a \mathbf{u} y \mathbf{v} se tiene que

$$\frac{\partial E}{\partial \mathbf{u}} = 2\tilde{\mathbf{D}}_{11} \mathbf{u} + 2\tilde{\mathbf{D}}_{12} \mathbf{v} - 2\lambda \mathbf{S}_b \mathbf{u} \quad (2.12)$$

$$\frac{\partial E}{\partial \mathbf{v}} = 2\tilde{\mathbf{D}}_{12}^T \mathbf{u} + 2\tilde{\mathbf{D}}_{22} \mathbf{v} \quad (2.13)$$

de donde

$$(\tilde{\mathbf{D}}_{11} - \lambda \mathbf{S}_b) \mathbf{u} + \tilde{\mathbf{D}}_{12} \mathbf{v} = 0 \quad (2.14)$$

$$\mathbf{v} = -\tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T \mathbf{u} \quad (2.15)$$

substituyendo el valor de \mathbf{v} en (2.14) se obtiene que

$$\begin{aligned} (\tilde{\mathbf{D}}_{11} - \lambda \mathbf{S}_b) \mathbf{u} - \tilde{\mathbf{D}}_{12} \tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T \mathbf{u} &= 0 \\ (\tilde{\mathbf{D}}_{11} - \tilde{\mathbf{D}}_{12} \tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T) \mathbf{u} &= \lambda \mathbf{S}_b \mathbf{u} \\ \mathbf{W} \mathbf{u} &= \lambda \mathbf{S}_b \mathbf{u} \end{aligned}$$

entonces λ es un eigenvalor generalizado. Como \mathbf{S}_b es invertible se llega al problema de eigenvalores

$$\mathbf{S}_b^{-1} \mathbf{W} \mathbf{u} = \lambda \mathbf{u} \quad (2.16)$$

Si se toma \mathbf{u}_1 asociado al eigenvalor más pequeño de (2.16), y se substituye $\mu \mathbf{u}_1$ en (2.11) se obtiene que

$$\mu \mathbf{u}_1^T \mathbf{S}_b \mu \mathbf{u}_1 = 1 \quad (2.17)$$

$$\mu^2 = \frac{1}{\mathbf{u}_1^T \mathbf{S}_b \mathbf{u}_1} \quad (2.18)$$

$$\mu = \frac{1}{\sqrt{\mathbf{u}_1^T \mathbf{S}_b \mathbf{u}_1}} \quad (2.19)$$

ahora bien, nombrando $\hat{\mathbf{u}} = \mu \mathbf{u}_1$ la solución queda

$$\mathbf{a} = \begin{pmatrix} \hat{\mathbf{u}} \\ -\tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T \hat{\mathbf{u}} \end{pmatrix}.$$

Normalización con $4AC - B^2 = 1$

Esta normalización fue propuesta recientemente por Fitzgibbon [9]. Su principal característica es que incorpora la especificidad de elipse.

Hay que recordar que el discriminante indica que cónica está representada por los coeficientes de la ecuación general de segundo grado y en el caso de la elipse éste debe ser negativo. La restricción $B^2 - 4AC < 0$ resulta difícil de manejar en general. sin embargo, como se tiene la libertad de escalar los parámetros arbitrariamente, basta con pedir $4AC - B^2 = 1$.

Halír [17] estudia este método y propone una formulación alternativa basada en la descomposición de matrices por bloques que contribuye a que el

algoritmo sea numéricamente estable. El procedimiento que sigue este autor es muy similar al que se usó para la normalización $A^2 + B^2/2 + C^2 = 1$.

Considere el problema

$$\min \sum_{i=1}^n Q(\mathbf{a}; \mathbf{x}_i)^2 = \min \|\mathbf{D}_X \mathbf{a}\|^2$$

Sea \mathbf{S}_F la matriz

$$\mathbf{S}_F = \begin{pmatrix} \mathbf{S}_f & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{pmatrix} \quad \text{donde} \quad \mathbf{S}_f = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

separando a la matriz $\tilde{\mathbf{D}} = \mathbf{D}_X^T \mathbf{D}_X$ por bloques

$$\tilde{\mathbf{D}} = \begin{pmatrix} \tilde{\mathbf{D}}_{11} & \tilde{\mathbf{D}}_{12} \\ \tilde{\mathbf{D}}_{12}^T & \tilde{\mathbf{D}}_{22} \end{pmatrix}$$

y al vector $\mathbf{a}^T = (\mathbf{u}^T \ \mathbf{v}^T)$ donde $\mathbf{u} = (A, B, C)^T$ y $\mathbf{v} = (D, E, F)^T$, entonces el problema

$$\begin{aligned} \min \quad & \mathbf{a}^T \mathbf{D}_X^T \mathbf{D}_X \mathbf{a} \\ \text{s.a.} \quad & \mathbf{a}^T \mathbf{S}_F \mathbf{a} = 1 \end{aligned}$$

puede ser reescrito como

$$\min_{\mathbf{u}, \mathbf{v}} \quad \mathbf{u}^T \tilde{\mathbf{D}}_{11} \mathbf{u} + 2\mathbf{u}^T \tilde{\mathbf{D}}_{12} \mathbf{v} + \mathbf{v}^T \tilde{\mathbf{D}}_{22} \mathbf{v} \quad (2.20)$$

$$\text{s.a.} \quad \mathbf{u}^T \mathbf{S}_f \mathbf{u} = 1 \quad (2.21)$$

Usando los multiplicadores de Lagrange se llega al sistema de ecuaciones

$$W \mathbf{u} = \lambda \mathbf{S}_f \mathbf{u} \quad (2.22)$$

$$\mathbf{v} = -\tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T \mathbf{u} \quad (2.23)$$

donde $W = \tilde{\mathbf{D}}_{11} - \tilde{\mathbf{D}}_{12} \tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T$. Como la matriz \mathbf{S}_f es invertible la ecuación (2.22) conduce al problema de eigenvalores

$$\mathbf{S}_f^{-1} W \mathbf{u} = \lambda \mathbf{u} \quad (2.24)$$

Si se toma \mathbf{u}_1 asociado al eigenvalor positivo más pequeño de (2.24) y se substituye $\mu \mathbf{u}_1$ en (2.21) se tiene que

$$\mu \mathbf{u}_1^T \mathbf{S}_f \mu \mathbf{u}_1 = 1 \quad (2.25)$$

$$\mu^2 = \frac{1}{\mathbf{u}_1^T \mathbf{S}_f \mathbf{u}_1} \quad (2.26)$$

$$\mu = \frac{1}{\sqrt{\mathbf{u}_1^T \mathbf{S}_f \mathbf{u}_1}} \quad (2.27)$$

y con esto, nombrando $\hat{\mathbf{u}} = \mu \mathbf{u}_1$ la solución queda

$$\mathbf{a} = \begin{pmatrix} \hat{\mathbf{u}} \\ -\tilde{\mathbf{D}}_{22}^{-1} \tilde{\mathbf{D}}_{12}^T \hat{\mathbf{u}} \end{pmatrix}$$

Algunas observaciones sobre el ajuste por distancia algebraica

Una de las ventajas de usar el ajuste por distancia algebraica es que se pueden obtener soluciones satisfactorias a un bajo costo computacional, puesto que los métodos son directos. Además, la normalización del vector de parámetros aporta ciertas características a la solución, tales como la invarianza bajo transformaciones euclidianas o la especificidad de elipse.

Pero en general, no se puede garantizar que los resultados siempre sean satisfactorios. Existen al menos dos razones principales para sostener esto:

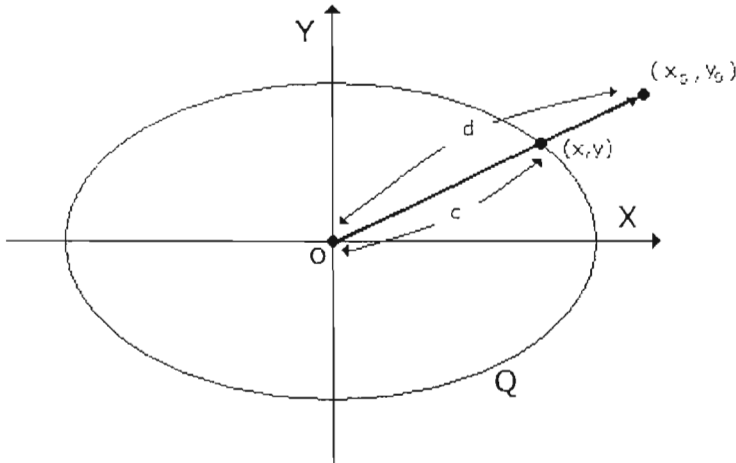
- i) La función a minimizar es generalmente no invariante bajo transformaciones euclidianas. Por ejemplo, la función con normalización $F = 1$ no es invariante bajo traslaciones.
- ii) Un punto puede contribuir a la estimación de los parámetros de manera diferente dependiendo de su posición en la cónica. Es decir, los métodos basados en las distancias algebraicas tienden a ajustar mejor la cónica a los puntos en las secciones de baja curvatura que en las secciones de alta curvatura. Este problema es usualmente llamado *sesgo por alta curvatura* y fue observado por Ballard [2].

Para ilustrar el problema del sesgo por alta curvatura considere la cónica

$$Q(x, y) = Ax^2 + Cy^2 + F = 0$$

con $AC > 0$; entonces la cónica corresponde a una elipse. Sea $\mathbf{p}_0 = (x_0, y_0)$ un punto fuera de la elipse y sea d la distancia del punto \mathbf{p}_0 al centro de la elipse. Se toma un punto $\mathbf{p} = (x, y)$ tal que esté sobre la elipse y que pertenezca al segmento que va del centro a \mathbf{p}_0 ; a la distancia de \mathbf{p} al centro se le llamará c . Por la colinealidad de \mathbf{p}_0 y \mathbf{p} se tiene que $x_0/y_0 = x/y$ de donde $x_0 = (x/y)y_0$ entonces

$$\begin{aligned} Q(x_0, y_0) &= Ax_0^2 + Cy_0^2 + F \\ &= A(x/y)^2 y_0^2 + Cy_0^2 + F \\ &= (y_0/y)^2 (Ax^2 + Cy^2) + F \\ &= (y_0/y)^2 (-F) + F \\ &= -F((y_0/y)^2 - 1) \end{aligned}$$



ya que (x, y) está sobre la elipse. Por semejanza de triángulos $y_0/y = d/c$, entonces

$$Q(x_0, y_0) = -F((y_0/y)^2 - 1) = -F((d/c)^2 - 1)$$

es decir que

$$Q(x_0, y_0) \propto \frac{d^2}{c^2} - 1$$

Si se mueve el punto p_0 a lo largo de la elipse de manera que su distancia a ésta siempre sea la misma, el cociente d/c no tendrá el mismo valor en todas las secciones de la elipse. Esto se debe a que el valor de c es mayor en una sección de alta curvatura que en una de baja curvatura. En consecuencia, el valor de la distancia de un punto a la elipse usando como métrica la distancia algebraica no es el mismo para un punto que está en una sección de alta curvatura que para un punto que esté en una sección de baja curvatura, aunque se encuentren a la misma distancia de la elipse con la métrica natural de \mathbb{R}^2 .

2.5.2 Ajuste por distancia geométrica

En esta sección se verá una forma de minimizar la distancia geométrica de los puntos a la elipse, para lo cual será necesario utilizar la representación paramétrica de la elipse.

Sea $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ un conjunto de puntos en \mathbb{R}^2 , ξ una representación de la elipse y $\mathcal{D}_G(\mathbf{X}, \xi)$ la distancia geométrica de los puntos a la elipse, entonces se quiere encontrar la elipse tal que:

$$\min_{\xi \in \text{elipse}} \mathcal{D}_G(\mathbf{X}, \xi)$$

Para abordar este problema se necesita de una representación que garantice que la solución corresponda a una elipse. En la sección anterior se vio que una forma de garantizar la especificidad de elipse usando como representación la ecuación general de segundo grado

$$Q(\mathbf{x}; \mathbf{a}) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$$

es restringiendo el espacio de parámetros de manera que se cumpla

$$4AC - B^2 = 1$$

Otra posibilidad, es representar a la elipse por medio de sus ecuaciones paramétricas como sugiere Gander [13]

$$\begin{aligned} x &= z_1 + a \cos \theta \\ y &= z_2 + b \sin \theta \end{aligned}$$

que en forma vectorial podemos escribir como

$$\mathbf{x} = \mathbf{z} + \begin{pmatrix} a \cos \theta \\ b \sin \theta \end{pmatrix}$$

donde $0 \leq \theta \leq 2\pi$. Como los ejes puedan ser cualesquiera, se usará la representación

$$\mathbf{x} = \mathbf{z} + Q(\alpha)\gamma(\theta) \begin{pmatrix} a \\ b \end{pmatrix}$$

donde $Q(\alpha)$ es la matriz de rotación y $\gamma(\theta)$ es una matriz diagonal; ambas se describen a continuación

$$\begin{aligned} Q(\alpha) &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \\ \gamma(\theta) &= \begin{pmatrix} \cos \theta & 0 \\ 0 & \sin \theta \end{pmatrix} \end{aligned}$$

Entonces en términos de la representación que se eligió la elipse queda caracterizada de la siguiente manera:

$$\xi(\alpha, a, b, \mathbf{z}) = \left\{ \mathbf{x}(\theta) = \mathbf{z} + Q(\alpha)\gamma(\theta) \begin{pmatrix} a \\ b \end{pmatrix} \mid 0 \leq \theta \leq 2\pi \right\}$$

con lo que la distancia de un punto a la cónica se puede definir como:

$$\begin{aligned} \mathcal{D}_G(\mathbf{x}_i; \xi(\alpha, a, b, \mathbf{z})) &= \min_{\mathbf{p} \in \xi} \|\mathbf{p} - \mathbf{x}_i\|^2 \\ &= \min_{\theta \in \mathbb{R}} \|\mathbf{p}(\theta) - \mathbf{x}_i\|^2 \end{aligned}$$

Entonces dado el conjunto de puntos \mathbf{X} se define la distancia de los puntos a la elipse como:

$$\begin{aligned} \mathcal{D}_G(\mathbf{X}, \xi) &= \sum_{i=1}^n \mathcal{D}_G(\mathbf{x}_i, \xi(\alpha, a, b, \mathbf{z})) \\ &= \min_{\theta \in \mathbb{R}} \|\mathbf{p}(\theta) - \mathbf{x}_1\|^2 + \dots + \min_{\theta \in \mathbb{R}} \|\mathbf{p}(\theta) - \mathbf{x}_n\|^2 \end{aligned}$$

con lo que nuestro problema se puede escribir como

$$\begin{aligned} \min_{\xi \in \text{elipse}} \mathcal{D}_G(\mathbf{X}, \xi) &= \min_{(\alpha, a, b, \mathbf{z})} \mathcal{D}_G(\mathbf{X}, \xi(\alpha, a, b, \mathbf{z})) \\ &= \min_{(\alpha, a, b, \mathbf{z})} \left(\sum_{i=1}^n \min_{\theta \in \mathbb{R}} \|\mathbf{p}(\theta) - \mathbf{x}_i\|^2 \right) \\ &= \min_{(\theta_1, \dots, \theta_n, \alpha, a, b, \mathbf{z})} \sum_{i=1}^n \|\mathbf{p}(\theta_i) - \mathbf{x}_i\|^2 \end{aligned}$$

Si se define una función \mathbf{g}_i de la siguiente manera

$$\mathbf{g}_i = \mathbf{x}_i - \mathbf{z} - Q(\alpha)\gamma(\theta_i) \begin{pmatrix} a \\ b \end{pmatrix}$$

entonces

$$\begin{aligned} \min_{(\alpha, a, b, \mathbf{z})} \mathcal{D}_G(\mathbf{X}, \xi(\alpha, a, b, \mathbf{z})) &= \min_{(\theta_1, \dots, \theta_n, \alpha, a, b, \mathbf{z})} \sum_{i=1}^n \|\mathbf{g}_i(\theta_i, \alpha, a, b, \mathbf{z})\|^2 \\ &= \min_{(\theta_1, \dots, \theta_n, \alpha, a, b, \mathbf{z})} \|\mathbf{g}(\theta_1, \dots, \theta_n, \alpha, a, b, \mathbf{z})\|^2 \end{aligned}$$

donde \mathbf{g} es un vector de $2n \times 1$

$$\mathbf{g} = \begin{pmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_n \end{pmatrix}$$

Entonces se calculan las derivadas parciales de \mathcal{D}_G con respecto a θ_j , α , a , b y z_k donde $j = 1, \dots, n$ y $k = 1, 2$.

$$\begin{aligned}\frac{\partial \mathcal{D}_G}{\partial \theta_j} &= \sum_{i=1}^n 2\mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial \theta_j} \\ \frac{\partial \mathbf{g}_i}{\partial \theta_j} &= -\delta_{ij} Q(\alpha) \gamma'(\theta_i) \begin{pmatrix} a \\ b \end{pmatrix}\end{aligned}$$

donde

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

y

$$\gamma'(\theta_i) = \begin{pmatrix} -\text{sen } \theta_i & 0 \\ 0 & \text{cos } \theta_i \end{pmatrix}$$

$$\begin{aligned}\frac{\partial \mathcal{D}_G}{\partial \alpha} &= \sum_{i=1}^n 2\mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial \alpha} \\ \frac{\partial \mathbf{g}_i}{\partial \alpha} &= -Q'(\alpha) \gamma(\theta_i) \begin{pmatrix} a \\ b \end{pmatrix}\end{aligned}$$

donde

$$Q'(\alpha) = \begin{pmatrix} -\text{sen } \alpha & -\text{cos } \alpha \\ \text{cos } \alpha & -\text{sen } \alpha \end{pmatrix}$$

$$\begin{aligned}\frac{\partial \mathcal{D}_G}{\partial a} &= \sum_{i=1}^n 2\mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial a} \\ \frac{\partial \mathbf{g}_i}{\partial a} &= -Q(\alpha) \gamma(\theta_i) \begin{pmatrix} 1 \\ 0 \end{pmatrix}\end{aligned}$$

$$\begin{aligned}\frac{\partial \mathcal{D}_G}{\partial b} &= \sum_{i=1}^n 2\mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial b} \\ \frac{\partial \mathbf{g}_i}{\partial b} &= -Q(\alpha) \gamma(\theta_i) \begin{pmatrix} 0 \\ 1 \end{pmatrix}\end{aligned}$$

$$\frac{\partial \mathcal{D}_G}{\partial z_k} = \sum_{i=1}^n 2\mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial z_k} \quad k = 1, 2.$$

$$\frac{\partial \mathbf{g}_i}{\partial z_k} = - \begin{pmatrix} \delta_{1k} \\ \delta_{2k} \end{pmatrix} \quad k = 1, 2.$$

Entonces el gradiente de la función \mathcal{D}_G se puede expresar de la siguiente manera:

$$\begin{aligned}
 \nabla \mathcal{D}_G &= \left(\frac{\partial \mathcal{D}_G}{\partial \theta_1}, \dots, \frac{\partial \mathcal{D}_G}{\partial \theta_n}, \frac{\partial \mathcal{D}_G}{\partial \alpha}, \frac{\partial \mathcal{D}_G}{\partial a}, \frac{\partial \mathcal{D}_G}{\partial b}, \frac{\partial \mathcal{D}_G}{\partial z_1}, \frac{\partial \mathcal{D}_G}{\partial z_2} \right)^T \\
 &= 2 \left(\sum_{i=1}^n \mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial \theta_1}, \dots, \sum_{i=1}^n \mathbf{g}_i^T \frac{\partial \mathbf{g}_i}{\partial z_2} \right)^T \\
 &= 2 \sum_{i=1}^n \left(\frac{\partial \mathbf{g}_i^T}{\partial \theta_1} \mathbf{g}_i, \dots, \frac{\partial \mathbf{g}_i^T}{\partial z_2} \mathbf{g}_i \right)^T \\
 &= 2 \sum_{i=1}^n \begin{pmatrix} \frac{\partial \mathbf{g}_i^T}{\partial \theta_1} \\ \vdots \\ \frac{\partial \mathbf{g}_i^T}{\partial z_2} \end{pmatrix} \mathbf{g}_i \\
 &= 2 \sum_{i=1}^n J_i^T \mathbf{g}_i = 2 J_{\mathbf{g}}^T \mathbf{g}
 \end{aligned}$$

donde las J_i son matrices de $2 \times (n+5)$ con un sólo elemento distinto de $\mathbf{0}_{2 \times 1}$ en la i -ésima de las primeras n columnas

$$\begin{aligned}
 J_i &= \left(\mathbf{0}_{2 \times 1} \mid \dots \mid \mathbf{0}_{2 \times 1} \mid -Q(\alpha)\gamma'(\theta_i) \begin{pmatrix} a \\ b \end{pmatrix} \mid \mathbf{0}_{2 \times 1} \mid \dots \mid \mathbf{0}_{2 \times 1} \dots \right. \\
 &\quad \left. -Q'(\alpha)\gamma(\theta_i) \begin{pmatrix} a \\ b \end{pmatrix} \mid -Q(\alpha)\gamma(\theta_i) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid -Q(\alpha)\gamma(\theta_i) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mid \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right)
 \end{aligned}$$

y $J_{\mathbf{g}}$ es la matriz Jacobiana:

$$J_{\mathbf{g}} = \begin{pmatrix} -Q\gamma_1'(b) & \dots & 0 & -Q'\gamma_1(b) & -Q\gamma_1(1) & -Q\gamma_1(0) & \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\
 \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & \dots & -Q\gamma_n'(a) & -Q'\gamma_n(a) & -Q\gamma_n(1) & -Q\gamma_n(0) & \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{pmatrix}$$

donde se han usado las abreviaciones Q para $Q(\alpha)$ y γ_i para $\gamma(\theta_i)$.

Dado que Q es ortogonal, la matriz $U = -\text{diag}(Q(\alpha), \dots, Q(\alpha))$ de $2n \times 2n$ es ortogonal también, entonces

$$U^T J_{\mathbf{g}} = \begin{pmatrix} \gamma_1'(b) & \dots & 0 & \tilde{Q}\gamma_1(a) & \gamma_1(1) & \gamma_1(0) & -Q^T \begin{pmatrix} -1 \\ 0 \end{pmatrix} & -Q^T \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\
 \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0 & \dots & \gamma_n'(a) & \tilde{Q}\gamma_n(b) & \gamma_n(1) & \gamma_n(0) & -Q^T \begin{pmatrix} -1 \\ 0 \end{pmatrix} & -Q^T \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{pmatrix}$$

donde \tilde{Q} es la matriz

$$\tilde{Q} = Q^T Q' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Si se permutan las ecuaciones en la matriz anterior se obtiene la siguiente estructura para la matriz Jacobiana

$$\bar{J} = P(U^T J_g) = \begin{pmatrix} -aS & A \\ bC & B \end{pmatrix}$$

donde P es una matriz de permutación, $S = \text{diag}(\sin \theta_i)$ y $C = \text{diag}(\cos \theta_i)$ son dos matrices diagonales de $n \times n$ y A y B son de $n \times 5$ y están definidas por

$$\begin{aligned} A(i, 1:5) &= \begin{bmatrix} -b \sin \theta_i & \cos \theta & 0 & \cos \alpha & \sin \alpha \end{bmatrix} \\ B(i, 1:5) &= \begin{bmatrix} a \cos \theta_i & 0 & \sin \theta_i & -\sin \alpha & \cos \alpha \end{bmatrix} \end{aligned}$$

Entonces dado que

$$\min_{(\alpha, a, b, z)} \mathcal{D}_G(\mathbf{X}, \xi(\theta, \alpha, a, b, z)) = \min_{(\theta_1, \dots, \theta_n, \alpha, a, b, z)} \| \mathbf{g}(\theta_1, \dots, \theta_n, \alpha, a, b, z) \|^2$$

si se cuenta con una aproximación inicial $\mathbf{x}_0 = (\theta_1^0, \dots, \theta_n^0, \alpha^0, a^0, b^0, z^0)$ se tiene que

$$\min_{(\alpha, a, b, z)} \mathcal{D}_G(\mathbf{X}, \xi(\theta, \alpha, a, b, z)) \approx \| \mathbf{g}(\theta_1^0, \dots, \theta_n^0, \alpha^0, a^0, b^0, z^0) + \bar{J}^0 \mathbf{h} \|^2$$

donde

$$\mathbf{h} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \\ \alpha \\ a \\ b \\ z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} \theta_1^0 \\ \vdots \\ \theta_n^0 \\ \alpha^0 \\ a^0 \\ b^0 \\ z_1^0 \\ z_2^0 \end{pmatrix}$$

Esto da lugar al problema lineal

$$\min_{\mathbf{h}} \| \mathbf{g}^0 + \bar{J}^0 \mathbf{h} \|^2$$

donde las ecuaciones normales

$$\bar{J}^{0T} \bar{J}^0 \mathbf{h} = -\bar{J}^{0T} \mathbf{g}^0$$

se resuelven por factorización QR para encontrar $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{h}$ y así sucesivamente hasta que la norma del residual $\|\mathbf{h}\|^2 < \epsilon$ donde $\epsilon > 0$ es una tolerancia.

Se pueden obtener aproximaciones iniciales para α, a, b, z ajustando una elipse por alguno de los métodos que minimizan la distancia algebraica. Por ejemplo, se puede usar el método para el ajuste por distancia algebraica sujeto a la restricción $4AC - B^2 = 1$ tomando ventaja del hecho de que con éste siempre se obtienen elipses. Una vez calculados los coeficientes de la ecuación general de segundo grado, se aplican las fórmulas que aparecen a continuación para calcular los valores de α, a, b, z .

Entonces, dada la ecuación general de segundo grado

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (2.28)$$

se puede eliminar el término xy rotando los ejes con un ángulo α tal que

$$\alpha = \frac{1}{2} \tan^{-1} \left(\frac{B}{A - C} \right) \quad (2.29)$$

con lo que la ecuación (2.28) se convierte en:

$$A'x^2 + C'y^2 + D'x + E'y + F' = 0 \quad (2.30)$$

donde

$$\begin{aligned} A' &= A \cos^2(\alpha) + C \sin^2(\alpha) + B \cos(\alpha) \sin(\alpha) \\ C' &= A \sin^2(\alpha) + C \cos^2(\alpha) - B \cos(\alpha) \sin(\alpha) \\ D' &= D \cos(\alpha) + E \sin(\alpha) \\ E' &= -D \sin(\alpha) + E \cos(\alpha) \\ F' &= F \end{aligned}$$

si se completan cuadrados en (2.30) se obtiene el centro de la cónica transformada

$$\begin{aligned} z'_1 &= -\frac{D'}{2A'} \\ z'_2 &= -\frac{E'}{2C'} \end{aligned}$$

aplicando de nuevo una rotación para volver a los ejes originales se obtiene el centro de cónica (2.28)

$$z_1 = z'_1 \cos(\alpha) - z'_2 \sin(\alpha) \quad (2.31)$$

$$z_2 = z'_1 \sin(\alpha) + z'_2 \cos(\alpha) \quad (2.32)$$

Para calcular la longitud de los semiejes considere las siguientes ecuaciones

$$k = F' - A'z_1'^2 - C'z_2'^2$$

$$a = \sqrt{\left| \frac{k}{A'} \right|} \quad (2.33)$$

$$b = \sqrt{\left| \frac{k}{C'} \right|} \quad (2.34)$$

Las aproximaciones iniciales para los ángulos $\theta_1, \dots, \theta_n$ se pueden obtener usando la siguiente fórmula

$$\theta_j = \arg \left(\frac{x_j'}{a} + i \frac{y_j'}{b} \right) \quad j = 1, \dots, n$$

donde i es el elemento unitario de los complejos, \arg es el argumento de un número complejo y x_j' y y_j' resultan de la transformación

$$\begin{pmatrix} x_j' \\ y_j' \end{pmatrix} = Q(\alpha) \begin{pmatrix} x_j - z_1 \\ y_j - z_2 \end{pmatrix}.$$

Capítulo 3

Resultados experimentales

3.1 Introducción

En este capítulo se examinan de forma experimental los métodos de ajuste de elipses que se revisaron en el capítulo anterior. Se comparan las soluciones que se obtienen de cada método bajo los siguientes experimentos: traslación o rotación de los puntos, adición de ruido a los puntos, ajuste con puntos que pasan por una hipérbola y ajuste con puntos que pasan por una elipse con alta excentricidad. Con el objeto de facilitar la realización de los experimentos se desarrolló un sistema¹ en MATLAB [22] con el que se pueden ajustar elipses usando cada uno de los métodos que se han revisado. El lector puede repetir cada uno los experimentos haciendo uso del sistema y cargando los puntos que se encuentra en el archivo que se indica.

De aquí en adelante, para facilitar la identificación de los métodos se hará la siguiente convención:

- *Paton*: ajuste por distancia algebraica con restricción $\| \mathbf{a} \|^2 = 1$.
- *Rosin*: ajuste por distancia algebraica con restricción $F = 1$.
- *Gander*: ajuste por distancia algebraica con restricción $A + C = 1$.
- *Bookstein*: ajuste por distancia algebraica con restricción $A^2 + B^2/2 + C^2 = 1$
- *Fitzgibbon*: ajuste por distancia algebraica con restricción $4AC - B^2 = 1$

¹Ver Apéndice C

- *Geométrico*: ajuste por distancia geométrica cuando la elipse se encuentra definida en forma paramétrica.

3.2 Comparación de los métodos

En los experimentos que se llevarán a cabo, se medirán algunas características geométricas en las elipses que se obtengan de aplicar cada uno de los métodos; las características geométricas que se van a medir son: las longitudes de los semiejes y la excentricidad de la elipse.

- *Invarianza*. Se compararán las soluciones que se obtiene al aplicar una transformación rígida a los puntos de prueba para cada uno de los métodos. El objetivo es observar si las longitudes de los semiejes de la elipse cambian.
- *Especificidad*. Se examinarán las soluciones que se obtienen con puntos de prueba que no están sobre una elipse.
- *Excentricidad*. Se ajustarán puntos de prueba que siguen la forma de una elipse con alta excentricidad para observar si se presentan diferencias en las longitudes de los ejes.
- *Ruido*. Se adicionará ruido gaussiano para observar la estabilidad del método así como las variaciones que se presenten en las características geométricas de la elipse.

Las cantidades que se presentan en las tablas de resultados fueron truncadas hasta la cuarta cifra decimal.

3.2.1 Invarianza

En la sección (1.5) se probó que las cantidades $A + C$, $B^2 - 4AC$ y $A^2 + B^2/2 + C^2$ son invariantes bajo rotaciones, pero que F y $\| \mathbf{a} \|^2$ no presentan esta propiedad. Fitzgibbon [10] demuestra la invarianza del discriminante para transformaciones afines y la invarianza de $A^2 + B^2/2 + C^2$ para transformaciones de similitud.

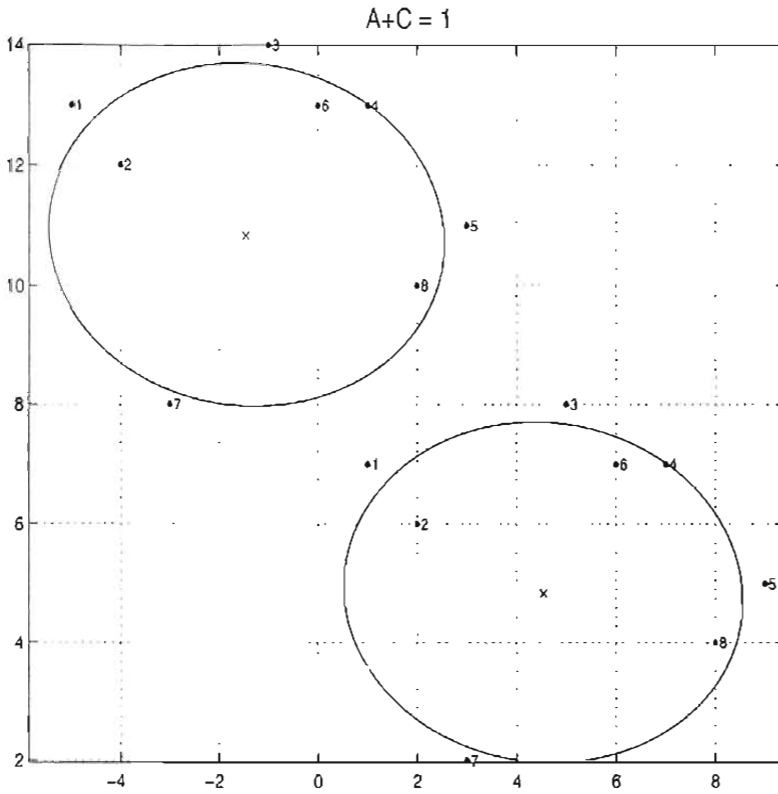
En los experimentos de esta sección se ajustará una elipse a unos puntos de prueba para comparar las características geométricas de la solución que se obtiene con las de la elipse que se ajusta a los puntos de prueba trasladados.

Considere el experimento de ajustar una elipse los puntos de la tabla (3.1) para compararla con la elipse que se ajusta al trasladar el origen a la coordenada $(-6, 6)$. Estos puntos de prueba fueron tomados de un experimento reportado por Gander [13] y se encuentran en el archivo `ptsinv1.txt`.

A la elipse que se ajusta a los puntos de prueba se le llamará de "referencia", mientras que a la que se obtiene de trasladar los puntos se le llamará "trasladada".

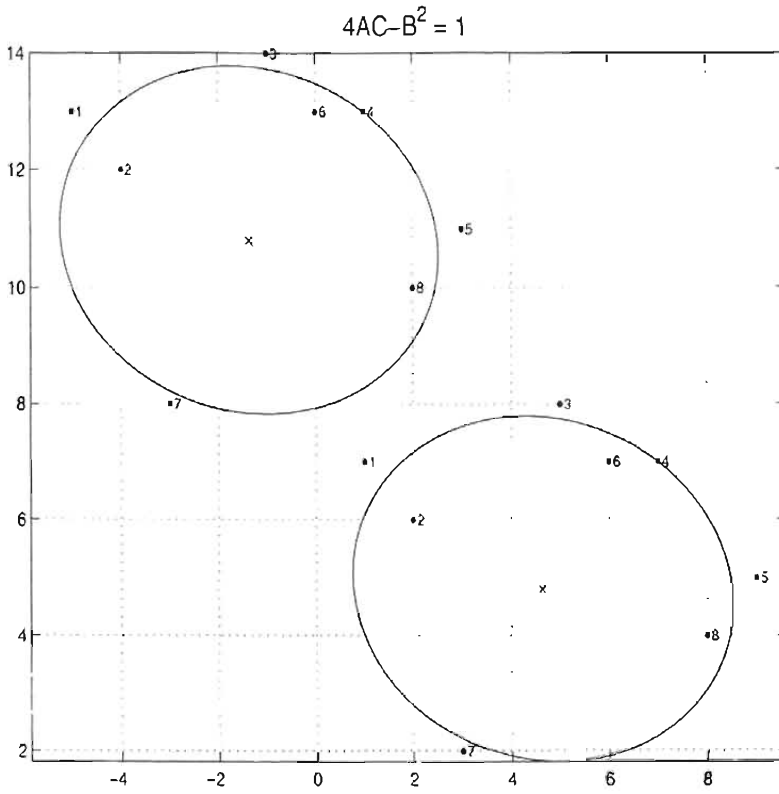
n°	x	y
1	1	7
2	2	6
3	5	8
4	7	7
5	9	5
6	6	7
7	3	2
8	8	4

Tabla 3.1: Colección de puntos de prueba



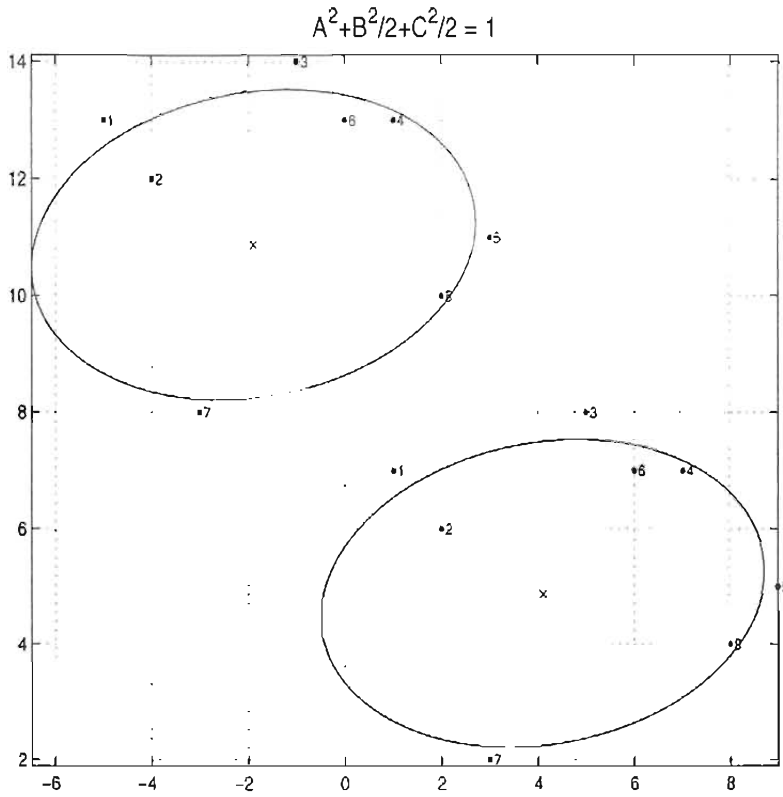
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(4.5384,4.8348)	4.0217	2.8634	0.7022
Trasladada	(-1.4615,10.8348)	4.0217	2.8634	0.7022

Figura 3.1: Resultados usando el método *Gander*



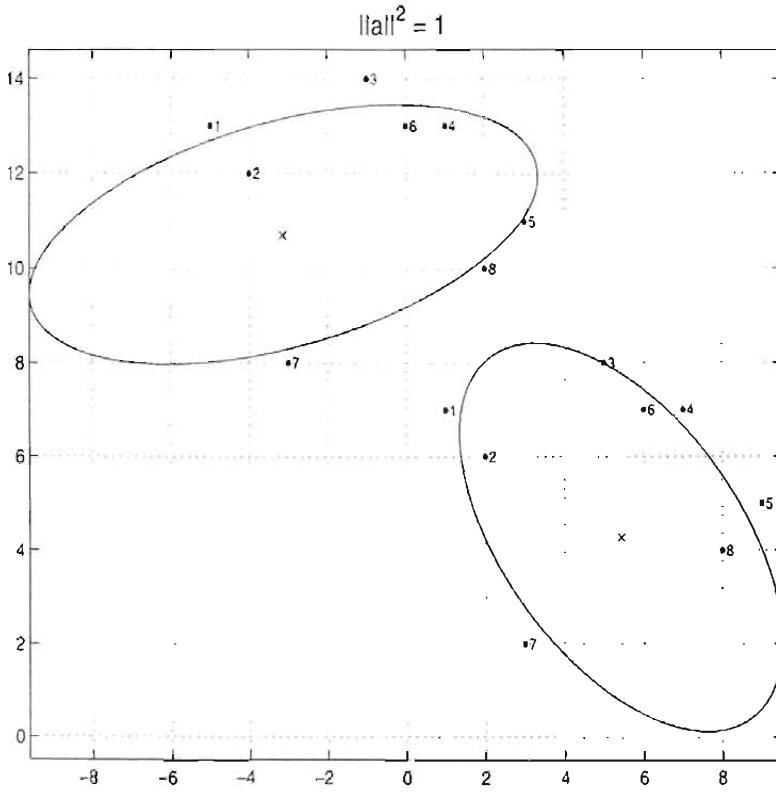
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(4.6360, 4.7984)	3.9099	2.9559	0.6546
Trasladada	(-1.3639, 10.7984)	3.9099	2.9559	0.6546

Figura 3.2: Resultados usando el método *Fitzgibbon*



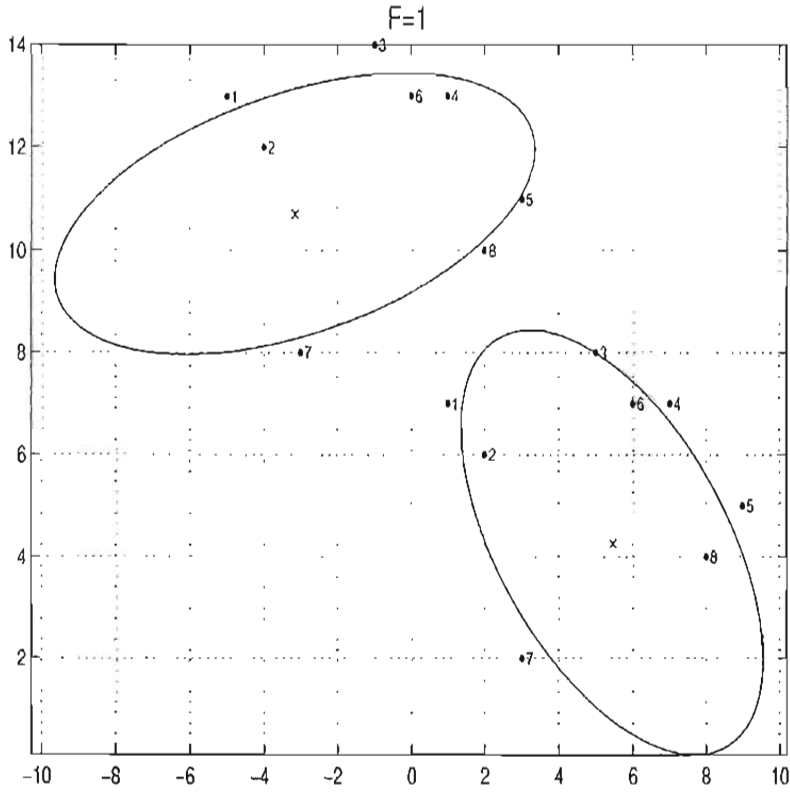
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(4.1092,4.8723)	4.6306	2.6141	0.8254
Trasladada	(-1.8907,10.8723)	4.6306	2.6141	0.8254

Figura 3.3: Resultados usando el método *Bookstein*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(5.4378,4.2705)	5.0971	2.8003	0.8356
Trasladada	(-3.1364,10.7075)	6.6158	2.3878	0.9326

Figura 3.4: Resultados usando el método *Paton*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(5.4593,4.2604)	5.1425	2.7886	0.8402
Trasladada	(-3.1480,10.7061)	6.6343	2.3865	0.9331

Figura 3.5: Resultados usando el método *Rosin*

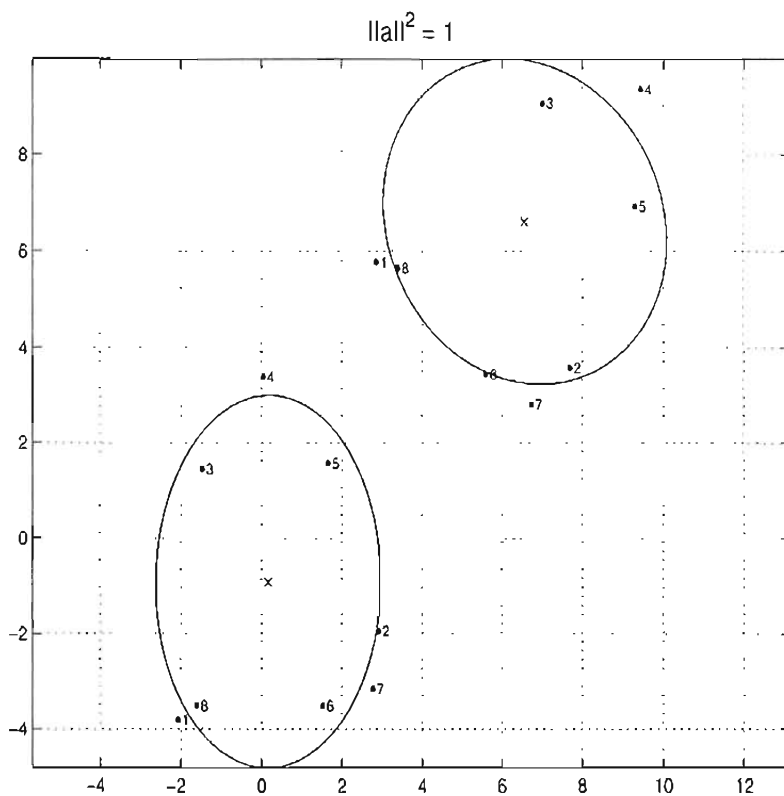
Como se puede ver en las gráficas y tablas anteriores los métodos *Gander*, *Fitzgibbon* y *Boosktein* encuentran la misma elipse después de trasladar los puntos. Por otra parte, en los métodos *Paton* y *Rosin* la elipse para los puntos trasladados no presenta la misma longitud en los semiejes.

En el siguiente ejemplo considere los puntos de prueba de la tabla (3.2), los cuales se obtuvieron de forma experimental y se encuentran en el archivo `ptsinv2.txt`. La transformación que se aplica a los puntos de prueba de este ejemplo consiste en trasladar el origen al punto $(7, 7)$ y rotar los ejes en un ángulo de $\pi/4$.

nº	x	y
1	-2.073733	-3.801170
2	2.903226	-1.929825
3	-1.474654	1.461988
4	0.046083	3.391813
5	1.658986	1.578947
6	1.520737	-3.508772
7	2.764977	-3.157895
8	-1.612903	-3.508772

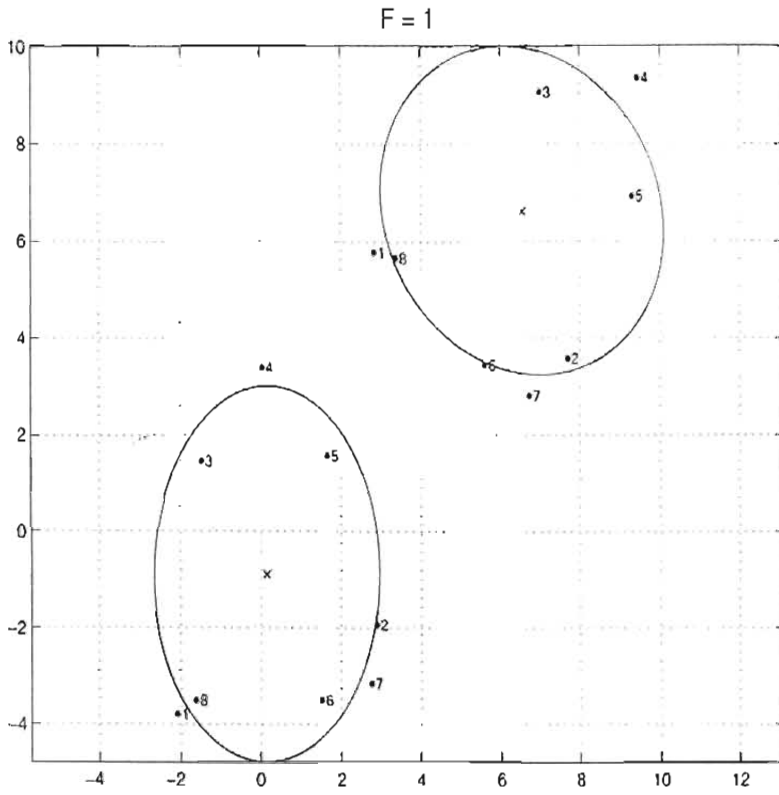
Tabla 3.2: Colección de puntos de prueba

En las gráficas y tablas que aparecen a continuación de nueva cuenta se observa que con los métodos *Paton* y *Rosin* se obtienen elipses con características geométricas distintas cuando los puntos son transformados. Por otra parte, con el método *Fitzgibbon* las elipses no cambian, tal como se esperaba. No se incluyen las gráficas correspondientes a los métodos *Gander* y *Bookstein* pero como era de esperarse, las elipses no cambian.



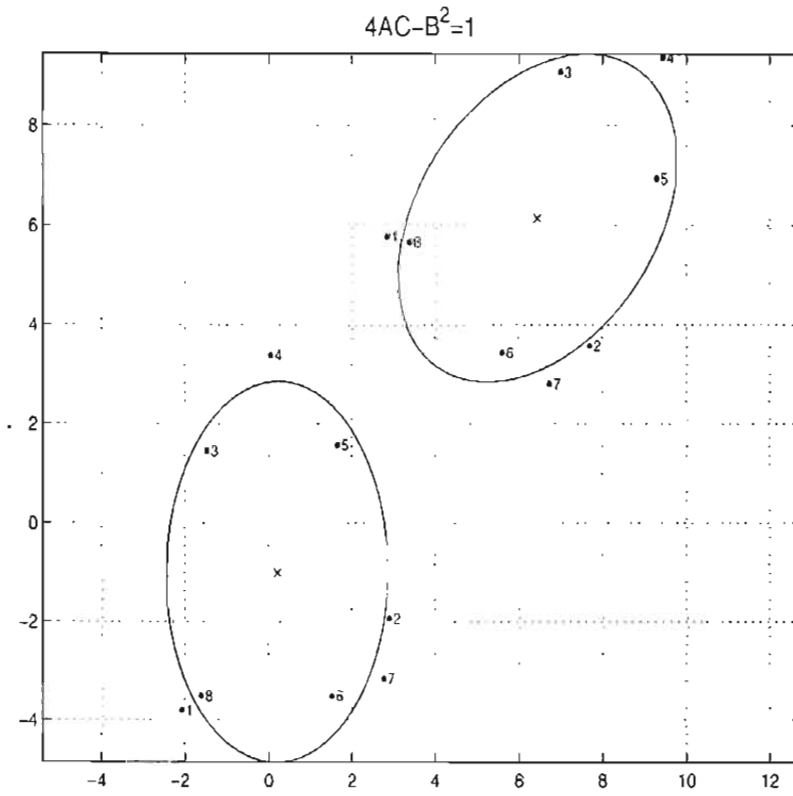
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0.1668,-0.9079)	3.9124	2.7884	0.7015
Trasladada	(6.5474,6.6192)	3.6680	3.2267	0.4755

Figura 3.6: Resultados usando el método *Paton*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0.1562,-0.8896)	3.9021	2.8072	0.6946
Trasladada	(6.5463,6.6291)	3.6942	3.2195	0.4904

Figura 3.7: Resultados usando el método *Rosin*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0.2067,-1.0029)	3.8609	2.6575	0.7254
Trasladada	(6.4369,6.1445)	3.8609	2.6575	0.7254

Figura 3.8: Resultados usando el método *Fitzgibbon*

En conclusión, la normalización del vector de parámetros a por una cantidad invariante bajo transformaciones euclideanas, es una de las características más importantes para el ajuste de elipses por distancia algebraica ya que aporta la capacidad de trasladar o rotar los puntos sin que la solución se vea afectada por la posición de los mismos en el plano.

3.2.2 Especificidad

Como se menciona en la sección (2.5) algunos métodos algebraicos no garantizan que la solución corresponda a una elipse. Por ejemplo, con los métodos *Paton*, *Rosin*, *Gander* y *Bookstein* se pueden obtener hipérbolas o parábolas según sea la forma que siguen los puntos que se quieren ajustar.

Por otra parte, el método *Fitzgibbon*, el cual incorpora la especificidad de elipse, deja marginadas a todas las demás cónicas debido a que normaliza la cantidad $4AC - B^2$. Con el método *Geométrico* también se obtienen únicamente elipses, sin embargo este método no es directo y en algunos casos no converge.

En el siguiente experimento considere los puntos de prueba descritos en la tabla (3.3) los cuales siguen la forma de una hipérbola; estos puntos se obtuvieron de forma experimental y se encuentran en el archivo `ptsesp1.txt`.

n°	x	y	n°	x	y
1	0.299539	3.508772	11	6.889401	9.619883
2	0.852535	3.654971	12	6.889401	8.187135
3	1.612903	3.976608	13	7.119816	7.865497
4	2.096774	3.830409	14	7.534562	7.660819
5	2.880184	3.596491	15	7.857143	7.660819
6	3.502304	3.333333	16	8.294931	7.456140
7	3.294931	2.514620	17	8.732719	7.251462
8	3.087558	2.017544	18	9.009217	7.251462
9	2.857143	1.345029	19	9.516129	7.280702
10	2.857143	0.409357	20	7.235023	8.947368

Tabla 3.3: Puntos de prueba

Las gráficas de los ajustes por los seis métodos se muestran en la siguiente página.

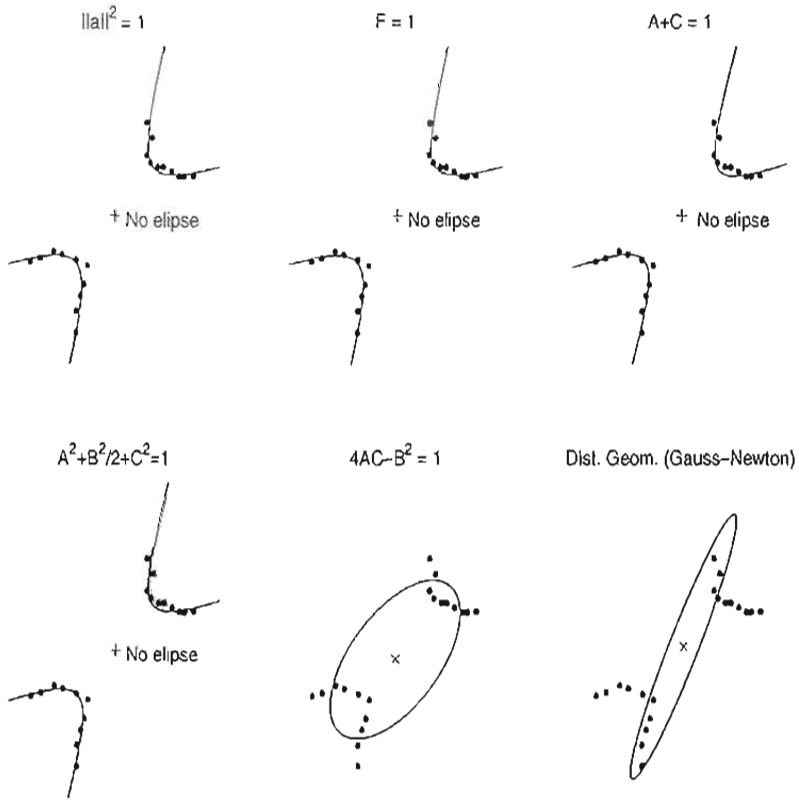


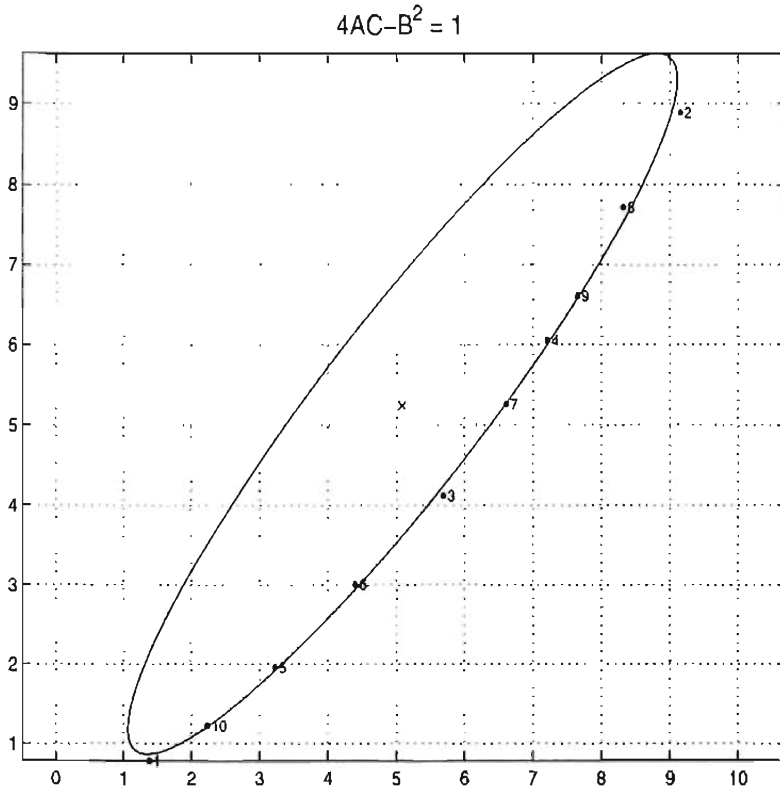
Figura 3.9: Cónicas que se obtienen para los puntos de la tabla (3.3)

Los métodos *Paton*, *Rosin*, *Gander* y *Bookstein* encuentran una hipérbola que describe satisfactoriamente la forma que siguen los puntos. Estos métodos resultan apropiados en este caso ya que pueden ajustar cualquier tipo de cónica. Por otra parte, los métodos *Fitzgibbon* y *Geométrico* tratan de ajustar una elipse aunque el resultado no es satisfactorio.

Un segundo experimento se hará con puntos de prueba que siguen la forma de una elipse alargada. Estos puntos se obtuvieron de forma experimental y se encuentran en el archivo `ptsesp2.txt`.

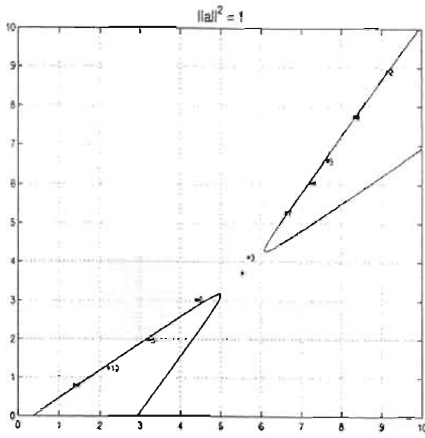
n°	x	y
1	1.382488	0.789474
2	9.147465	8.888889
3	5.691244	4.122807
4	7.211982	6.052632
5	3.225806	1.959064
6	4.400922	3.011696
7	6.612903	5.263158
8	8.317972	7.719298
9	7.649770	6.608187
10	2.235023	1.228070

Tabla 3.4: Puntos de prueba

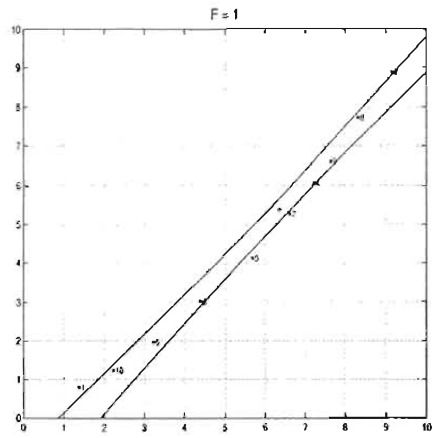


Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(5.0847, 5.2459)	5.8343	1.1245	0.9812

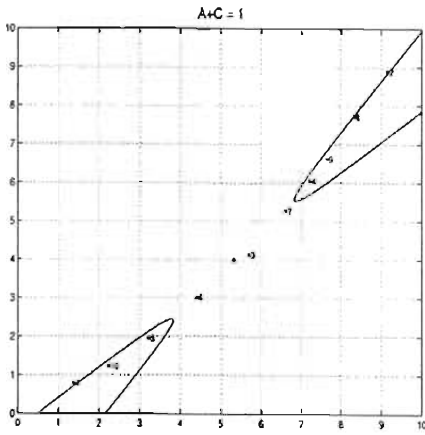
Figura 3.10: Resultados usando el método *Fitzgibbon*



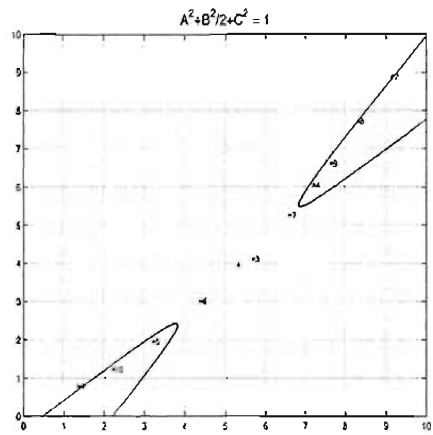
(a) Paton



(b) Rosin



(c) Gander



(d) Bookstein

Figura 3.11: Hipérbolas con los puntos de la tabla (3.4)

Después de hacer el ajuste de los puntos por cada método se obtuvieron los siguiente resultados: el método *Fitzgibbon* se ajusta razonablemente a los puntos siguiendo la forma elíptica de estos; por otra parte, los métodos *Paton*, *Rosin*, *Gander* y *Bookstein* encuentran hipérbolas. En el caso del método *Geométrico* después de 1000 iteraciones del método de Gauss-Newton no se llegó a una solución razonable.

3.2.3 Excentricidad

En muchos objetos de la naturaleza se encuentran formas similares a una elipse y en ocasiones, según sea nuestro ángulo de visión, se pueden observar elipses con alta excentricidad. Por ejemplo, en la imagen que aparece abajo se muestra un vaso cuyo borde corresponde a una elipse. Si se plantea el problema de identificar la elipse que aparece en el borde del vaso se debe tomar en cuenta que esta es considerablemente excéntrica.



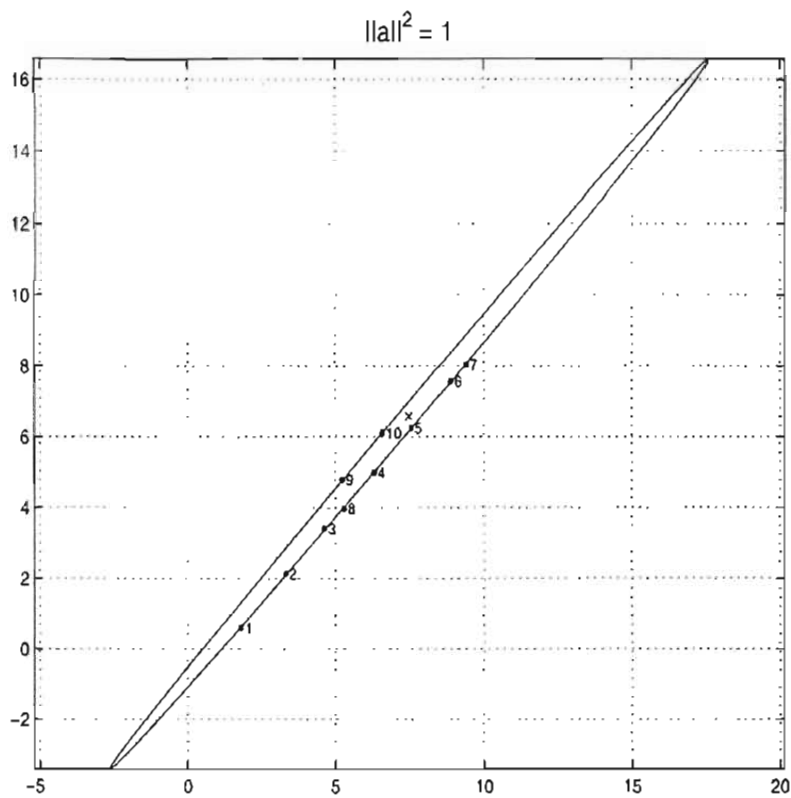
Figura 3.12: Elipse excéntrica en el borde de un vaso

En esta sección se va a experimentar con puntos de prueba que siguen la forma de una elipse alargada. Considere los puntos que se muestran en la tabla (3.5) los cuales fueron generados de forma experimental y se encuentran en el archivo `ptsexcen1.txt`.

n°	x	y
1	1.797235	0.614035
2	3.317972	2.134503
3	4.631336	3.421053
4	6.313364	5.000000
5	7.557604	6.257310
6	8.870968	7.573099
7	9.400922	8.040936
8	5.299539	3.976608
9	5.230415	4.795322
10	6.589862	6.111111

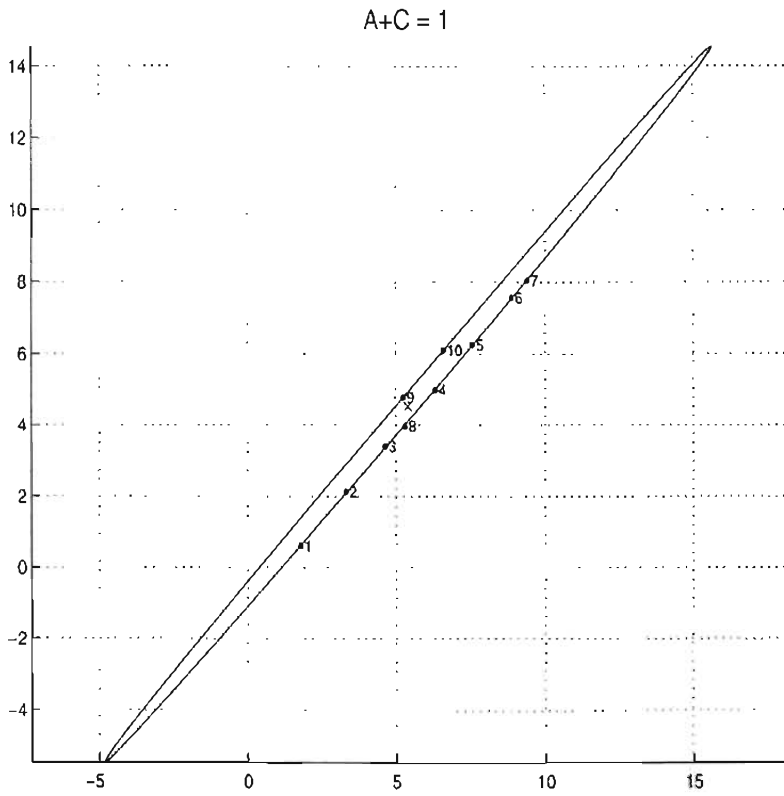
Tabla 3.5: Puntos de prueba

A continuación se presentan las gráficas de los ajustes por los métodos *Paton*, *Gander*, *Bookstein*, *Fitzgibbon* y *Geométrico* para estos puntos. Se omite el método *Rosin* ya que este encuentra una hipérbola.



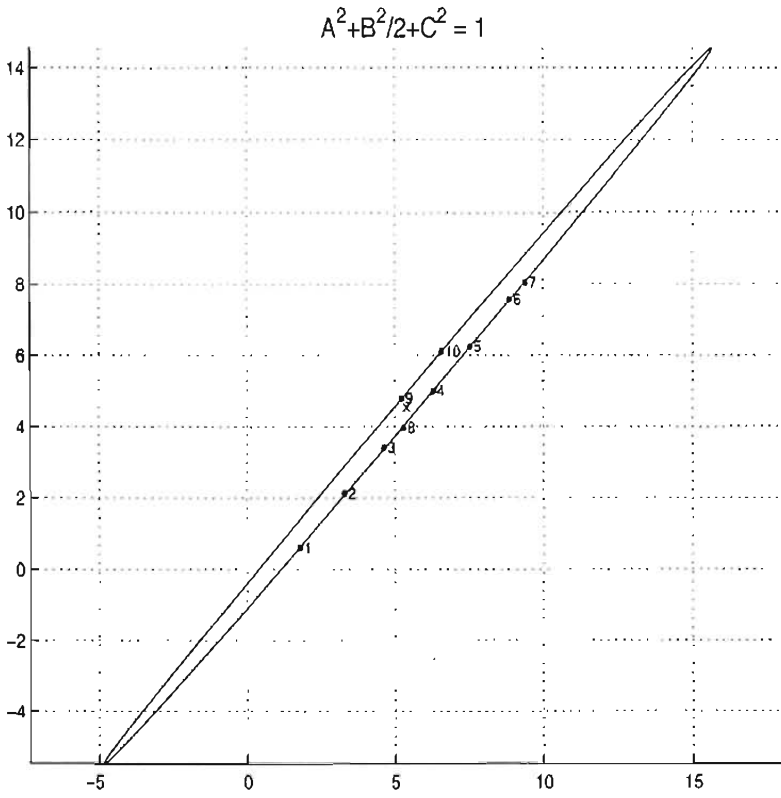
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(7.4833,6.5863)	14.2293	0.2976	0.9998

Figura 3.13: Resultados usando el método *Paton*



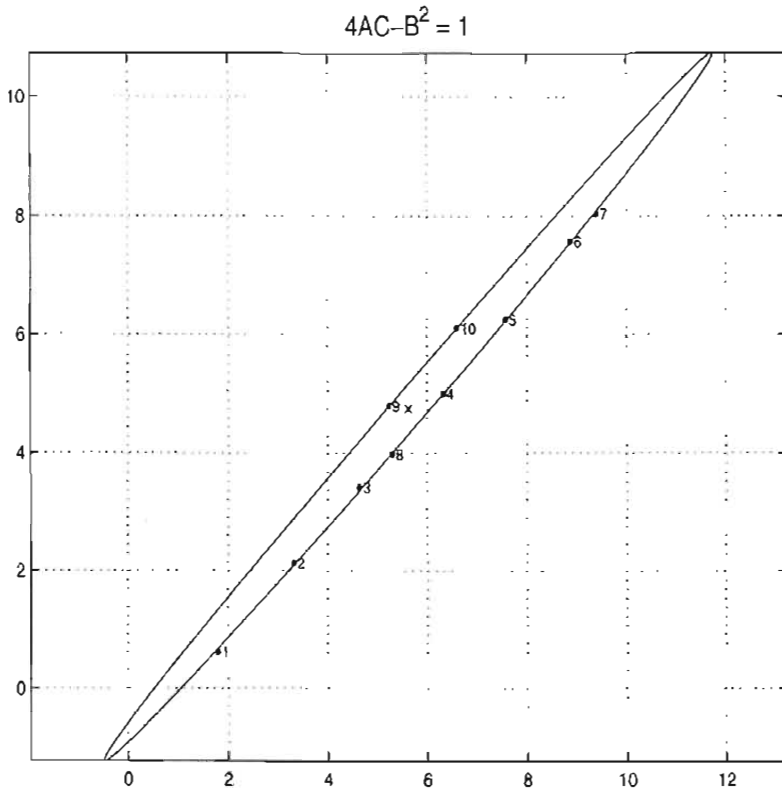
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(5.3999,4.5379)	14.3211	0.2957	0.9998

Figura 3.14: Resultados usando el método *Gander*



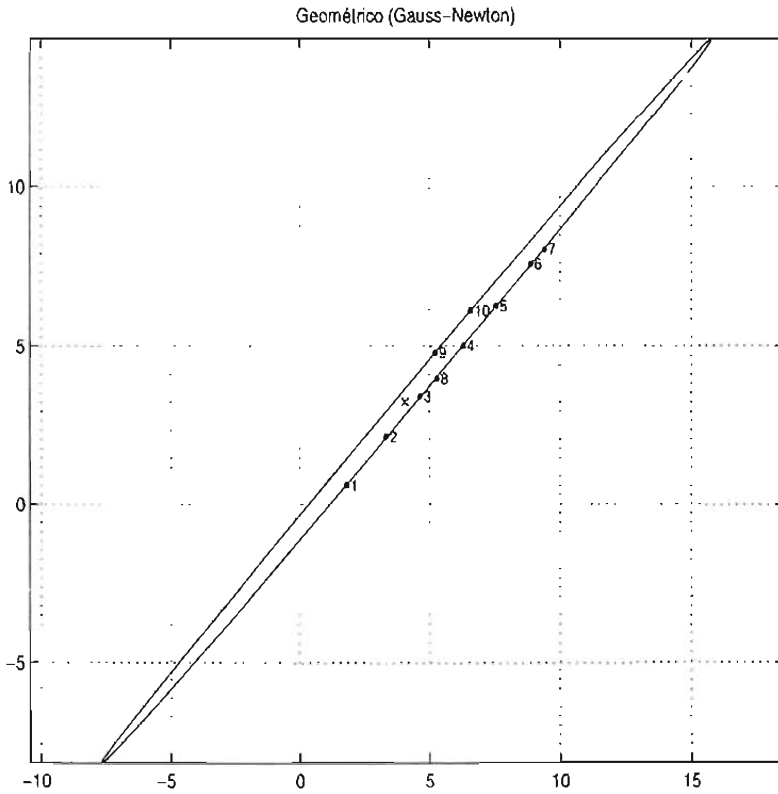
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(5.3983,4.5364)	14.3449	0.2957	0.9998

Figura 3.15: Resultados usando el método *Bookstein*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(5.6283, 4.7503)	8.5385	0.3053	0.9994

Figura 3.16: Resultados usando el método *Fitzgibbon*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(4.0581,3.2398)	16.3806	0.2995	0.9998

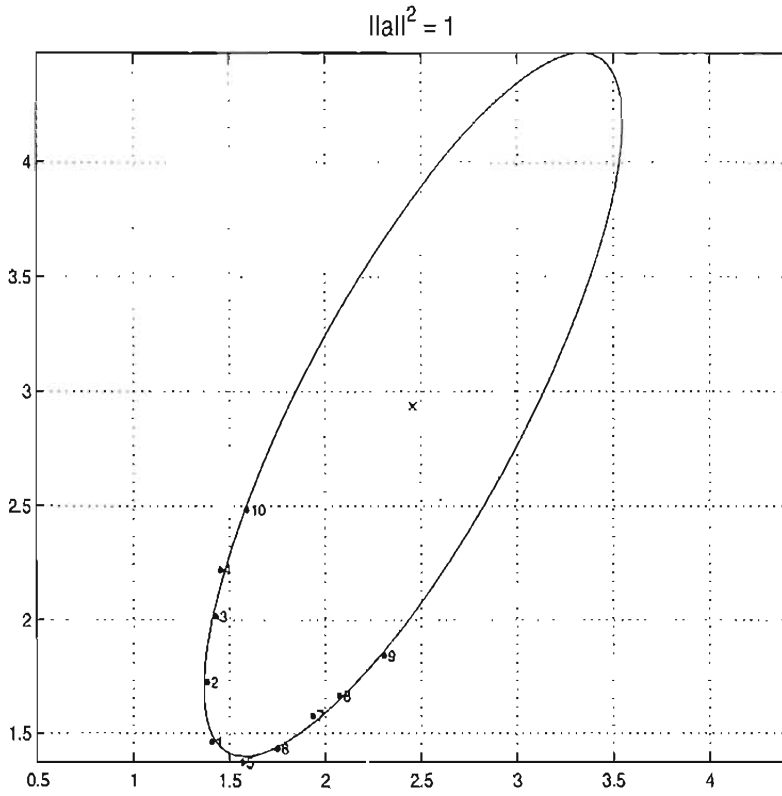
Figura 3.17: Resultados usando el método *Geométrico*

Como se pudo observar, la excentricidad de las elipses ajustadas por cada uno de los métodos no varía demasiado. Por otra parte, se encontró que la longitud del semieje mayor para la elipse del método *Fitzgibbon* fue considerablemente menor a las longitudes encontradas para las demás elipses.

Considere los puntos de prueba de la tabla (3.6) los cuales se eligieron experimentalmente de manera que se encuentren cerca de uno de los vértices, con la finalidad de analizar el comportamiento de los métodos en secciones de alta curvatura; estos se encuentran en el archivo ptsexcen2.txt.

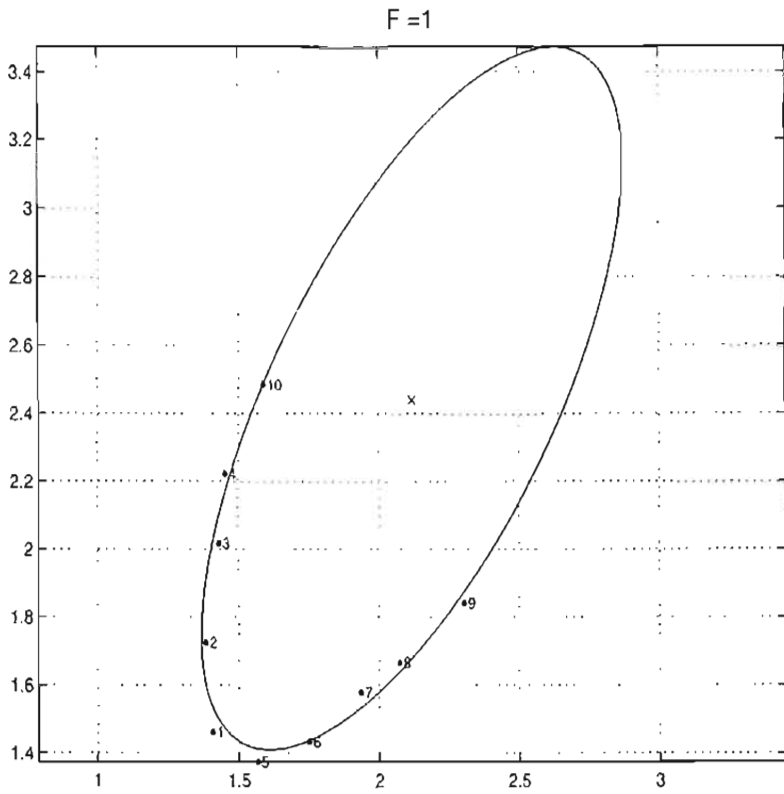
n ^o	x	y
1	1.405530	1.461988
2	1.382488	1.725146
3	1.428571	2.017544
4	1.451613	2.222222
5	1.566820	1.374269
6	1.751152	1.432749
7	1.935484	1.578947
8	2.073733	1.666667
9	2.304147	1.842105
10	1.589862	2.485380

Tabla 3.6: Puntos de prueba



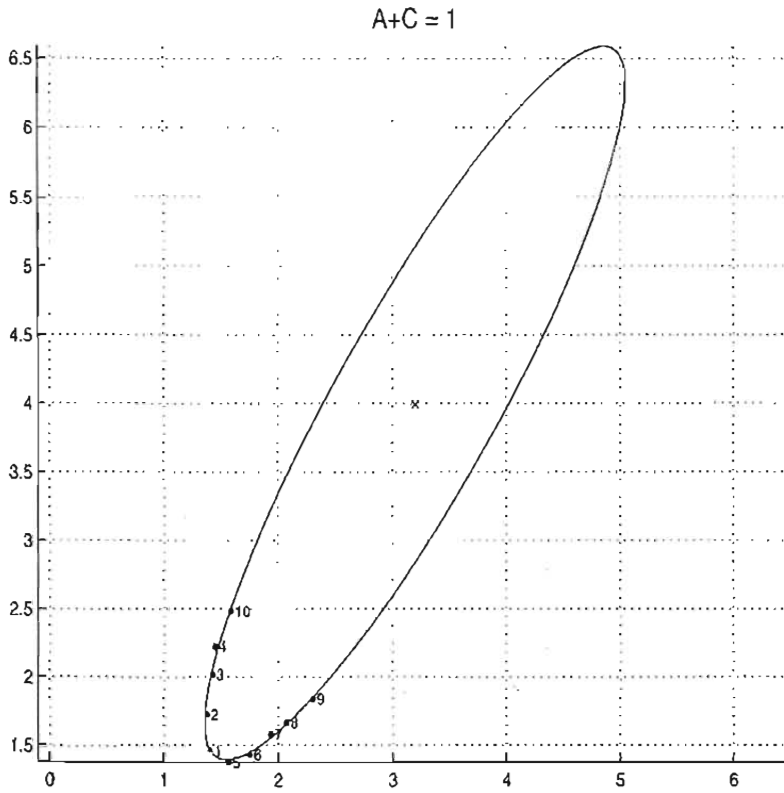
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(2.4559,2.9362)	1.8014	0.5503	0.9939

Figura 3.18: Resultados usando el método *Paton*



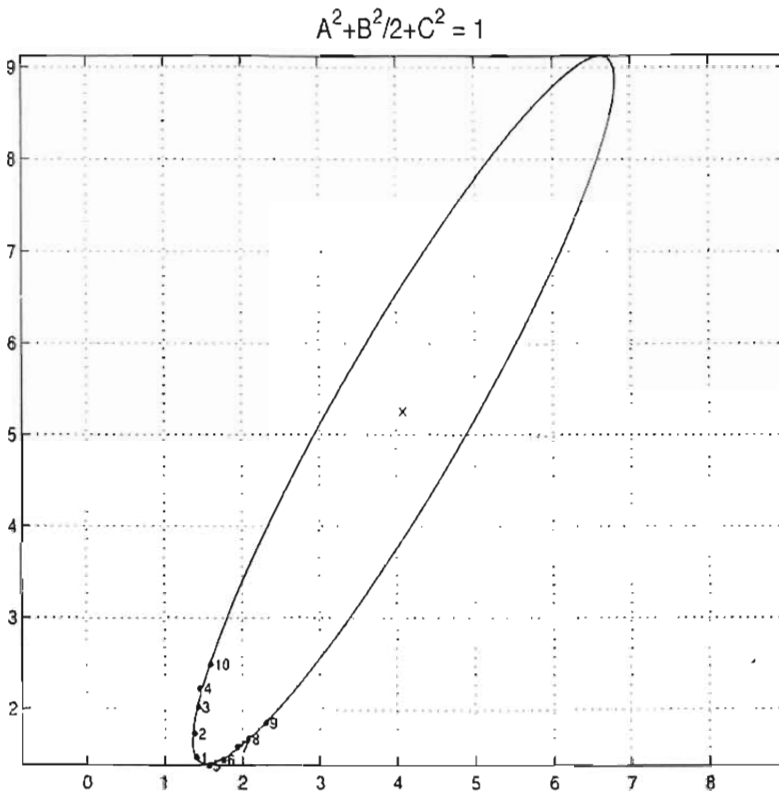
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(2.1194,2.4405)	1.1817	0.4836	0.9124

Figura 3.19: Resultados usando el método *Rosin*



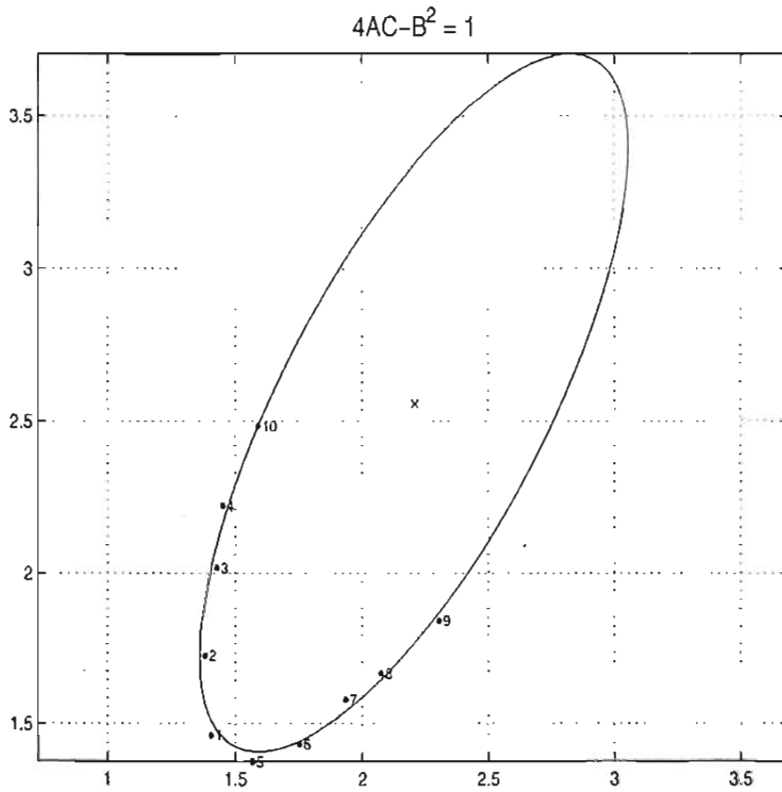
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(3.2024,3.9929)	3.1124	0.6803	0.9758

Figura 3.20: Resultados usando el método *Gander*



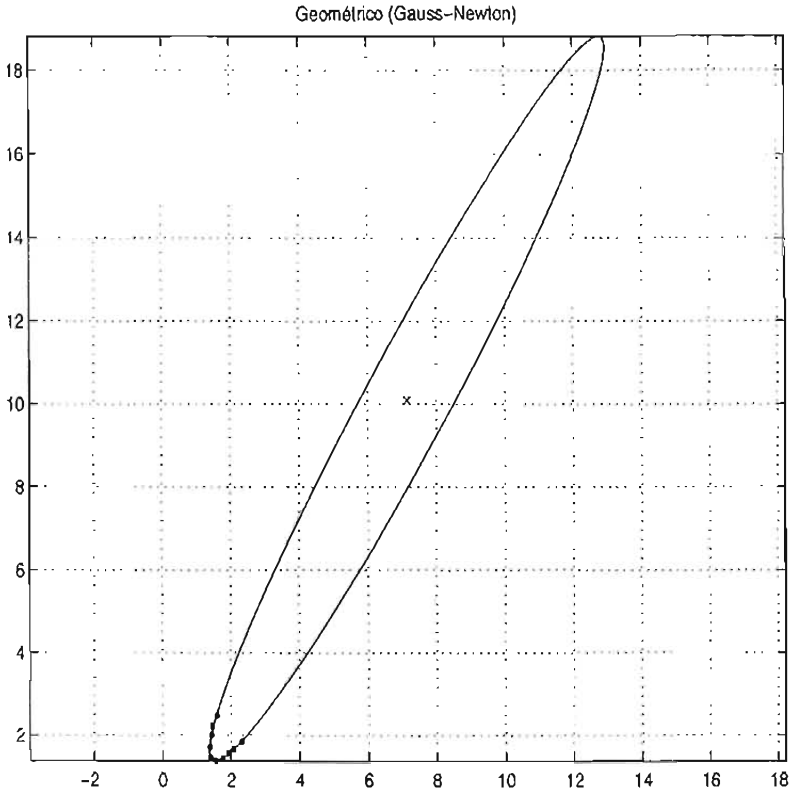
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(4.0790,5.2575)	4.6582	0.8104	0.9847

Figura 3.21: Resultados usando el método *Bookstein*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(2.2083, 2.5571)	1.3387	0.4992	0.9279

Figura 3.22: Resultados usando el método *Fitzgibbon*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
	(7.1552,10.0965)	10.3968	1.1734	0.9936

Figura 3.23: Resultados usando el método *Geométrico*

En este experimento se observa que con el método *Geométrico* se encuentra una elipse más alargada que las que se obtuvieron por los otros métodos y que los puntos se encuentran más cercanos a ésta; este comportamiento se debe a que este método minimiza la distancia geométrica. Es decir, se puede observar que minimizar la distancia algebraica de los puntos a la elipse en las secciones de alta curvatura puede no ser satisfactorio.

3.2.4 Ruido

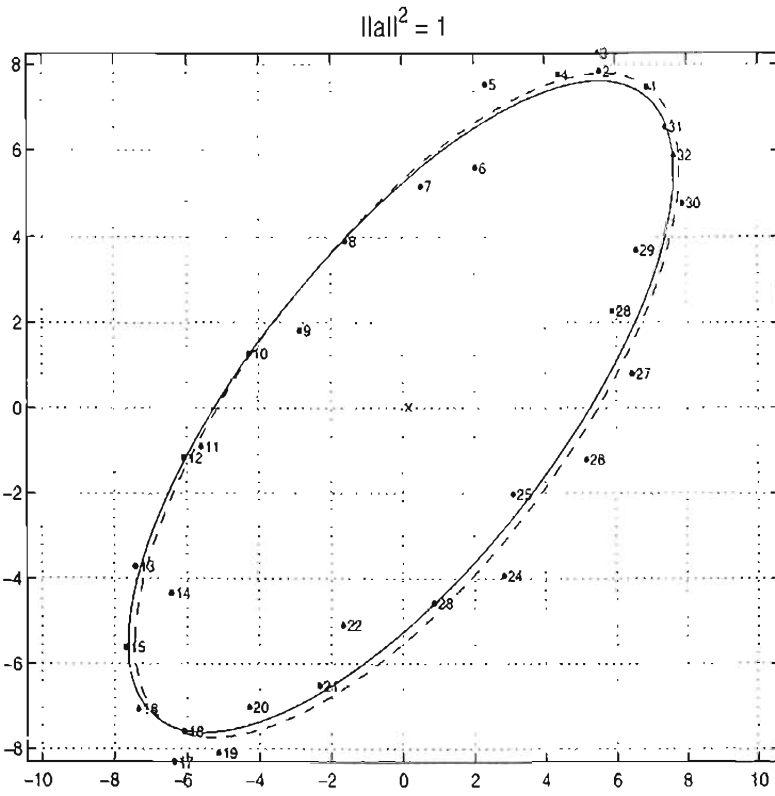
Por último se analizará el comportamiento de los métodos respecto a la presencia de ruido o perturbación de los datos. Para este experimento se tomaron 32 puntos sobre la elipse de referencia

$$29x^2 - 42xy + 29y^2 - 800 = 0$$

y fueron perturbados agregando ruido gaussiano en cada una de sus coordenadas. Los puntos que pertenecen a la elipse se encuentran en el archivo `ptsruido1a.txt` mientras que los puntos perturbados están en el archivo `ptsruido1b.txt`. A continuación se comparan la elipse de "referencia" (línea continua) con la "obtenida" al perturbar los puntos (línea discontinua) para cada uno de los métodos.

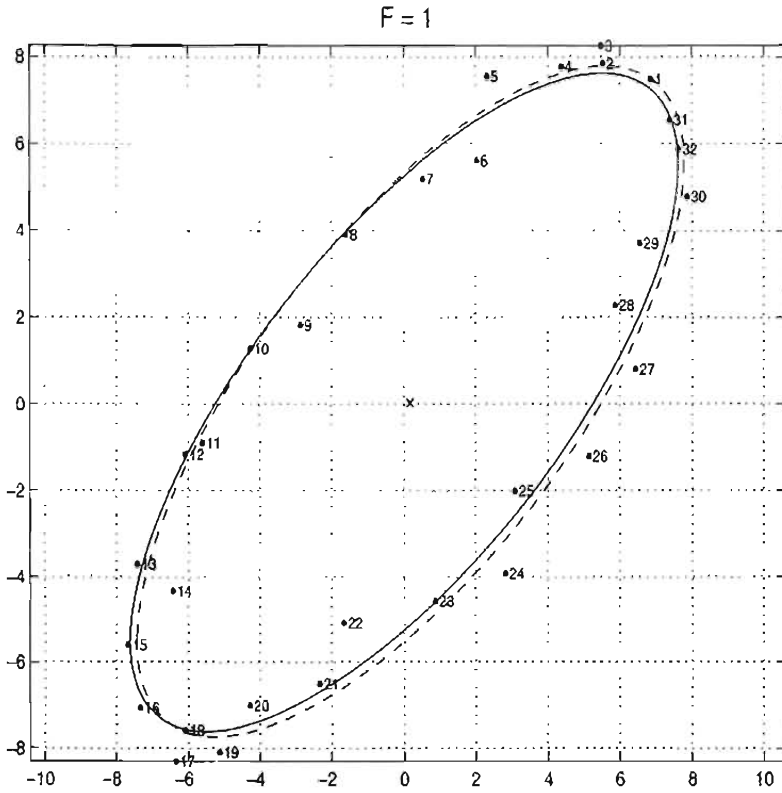
n°	x	y	n°	x	y
1	-6.893903	-7.224117	17	7.071068	7.071068
2	-6.113516	-7.559075	18	6.368196	7.492039
3	-5.089403	-7.592676	19	5.411443	7.614326
4	-3.862391	-7.323582	20	4.238954	7.433054
5	-2.481398	-6.762519	21	2.897471	6.955450
6	-1.001479	-5.931856	22	1.440475	6.200554
7	0.518366	-4.864709	23	-0.073948	5.198461
8	2.017544	-3.603621	24	-1.585424	3.989122
9	3.436290	-2.198868	25	-3.033693	2.620749
10	4.718042	-0.706453	26	-4.361018	1.147896
11	5.811700	0.814126	27	-5.514484	-0.370721
12	6.673664	2.302248	28	-6.448104	-1.874558
13	7.269571	3.698587	29	-7.124659	-3.303663
14	7.575662	4.947475	30	-7.517177	-4.601060
15	7.579736	5.999123	31	-7.610008	-5.715028
16	7.281629	6.811604	32	-7.399452	-6.601156

Tabla 3.7: Puntos en `ptsruido1a.m`



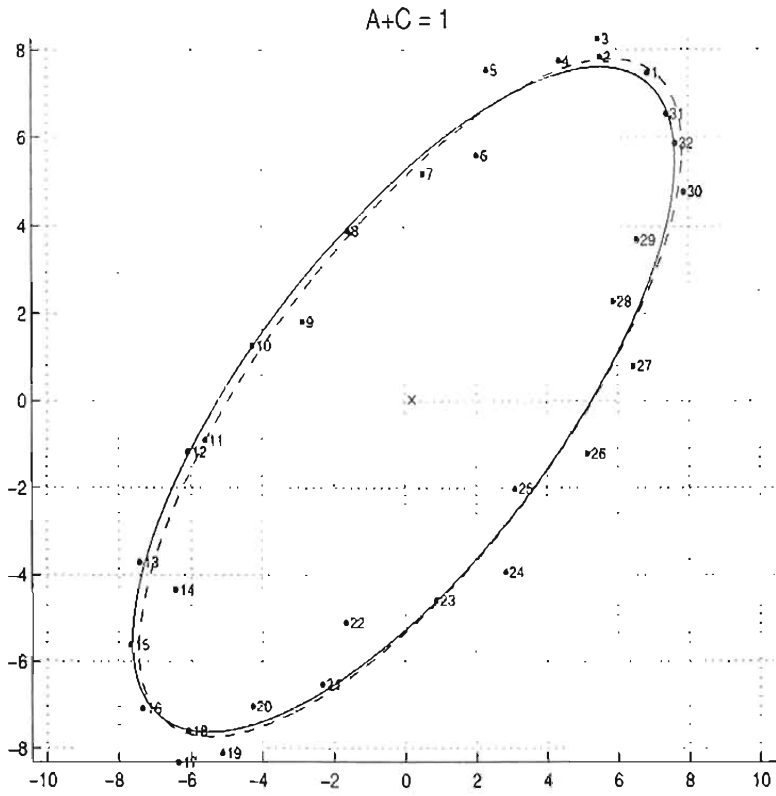
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	10	4	0.9165
Obtenida	(0.1729,0.0263)	10.0645	4.1033	0.9131

Figura 3.24: Resultados usando el método *Paton*



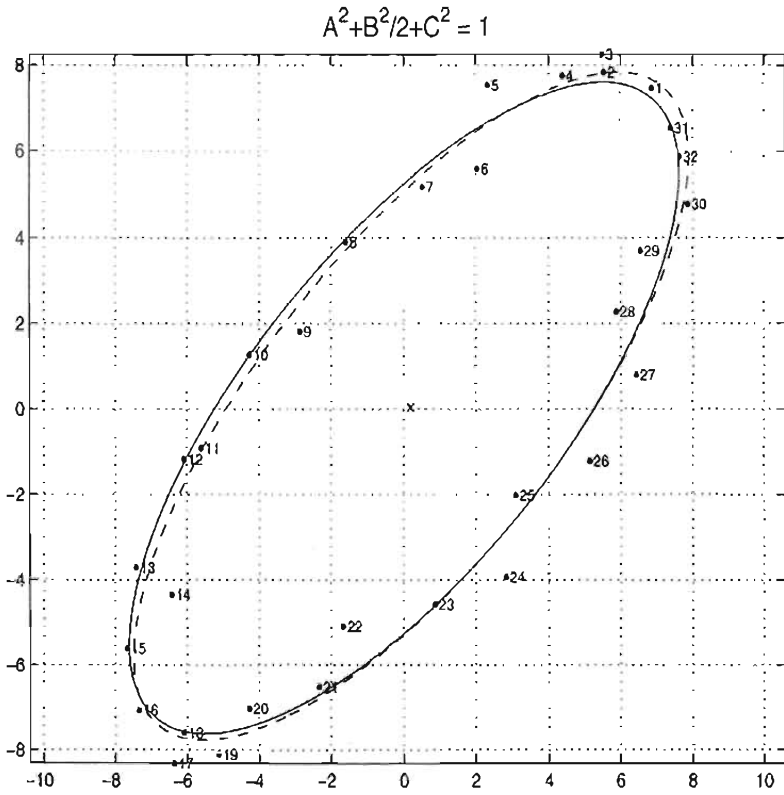
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	10	4	0.9165
Obtenida	(0.1724,0.0265)	10.0634	4.1046	0.9130

Figura 3.25: Resultados usando el método *Rosin*



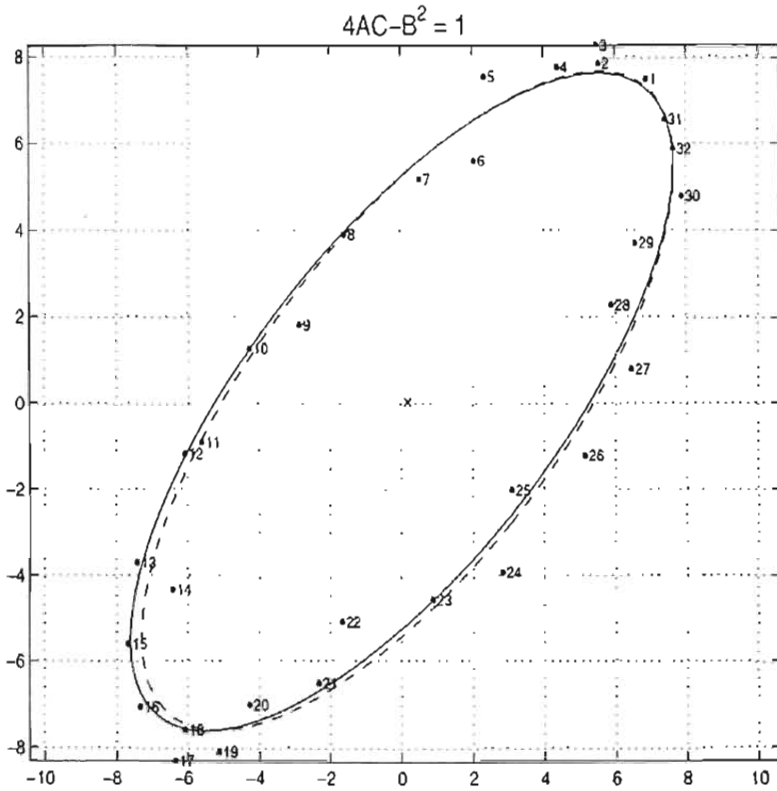
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	10	4	0.9165
Obtenida	(0.1897,0.0377)	10.1506	3.9206	0.9224

Figura 3.26: Resultados usando el método *Gander*



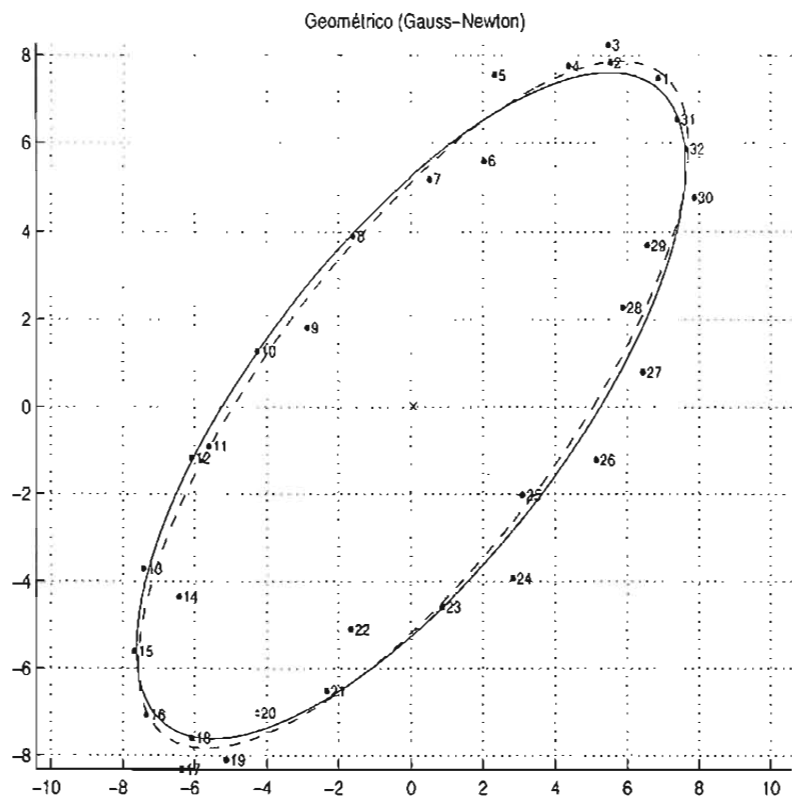
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	10	4	0.9165
Obtenida	(0.1934,0.0404)	10.2401	3.8878	0.9251

Figura 3.27: Resultados usando el método *Bookstein*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	10	4	0.9165
Obtenida	(0.1793, 0.0305)	9.8956	4.0243	0.9136

Figura 3.28: Resultados usando el método *Fitzgibbon*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	10	4	0.9165
Obtenida	(0.0647,0.0206)	10.2606	3.8276	0.9278

Figura 3.29: Resultados usando el método *Geométrico*

En este ejemplo se observa que los elementos que describen las características geométricas de la elipse no presentaron grandes desviaciones cuando se agregó ruido a los puntos. En éste y otros experimentos similares, se observó que la presencia de ruido no afecta demasiado las características geométricas de la elipse siempre y cuando los puntos se encuentren distribuidos a lo largo de toda ésta. Cuando los puntos solamente abarcan una región, las diferencias pueden ser más drásticas. Veamos un ejemplo.

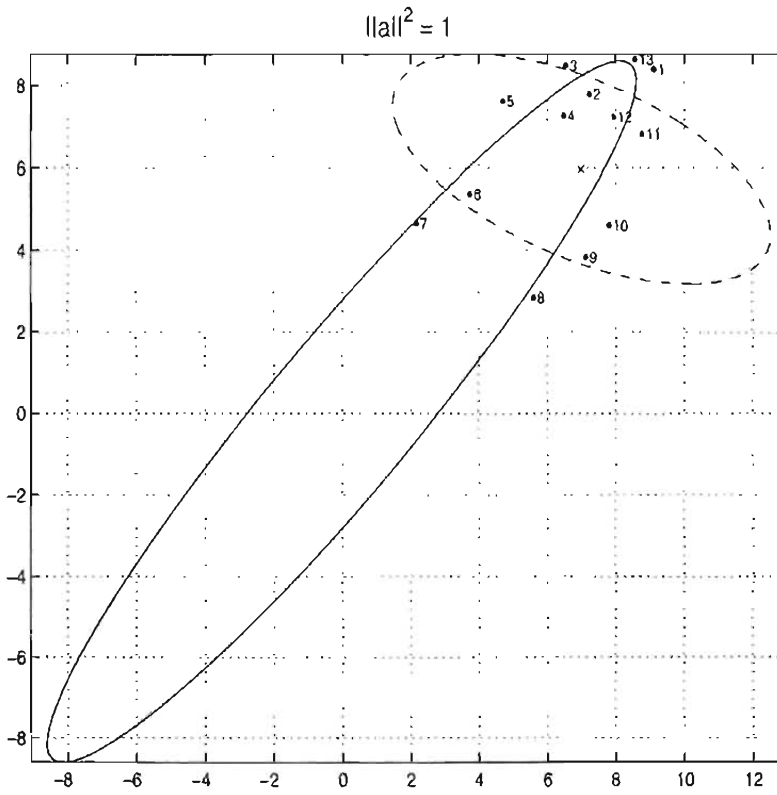
Considere los puntos de prueba que se muestran en la tabla (3.8). Estos puntos están distribuidos en la sección de alta curvatura de la elipse

$$74x^2 - 140xy + 74y^2 - 576 = 0$$

cuya excentricidad es $e = 0.986$. Con la finalidad de observar como afecta el ruido cuando los puntos se encuentran en una región de alta curvatura los puntos de prueba fueron perturbados agregando ruido gaussiano a cada una de sus coordenadas. Los puntos de prueba y los perturbados se encuentran en los archivos `ptsruido2a.txt` y `ptsruido2b.txt` respectivamente. Abajo se muestran los puntos originales y en las siguientes páginas se ven las gráficas de los ajustes para cada uno de los métodos; las elipses graficadas con línea continua se identificarán con el nombre de “referencia” y las elipse graficadas con línea discontinua se identificarán con el nombre de “obtenida”.

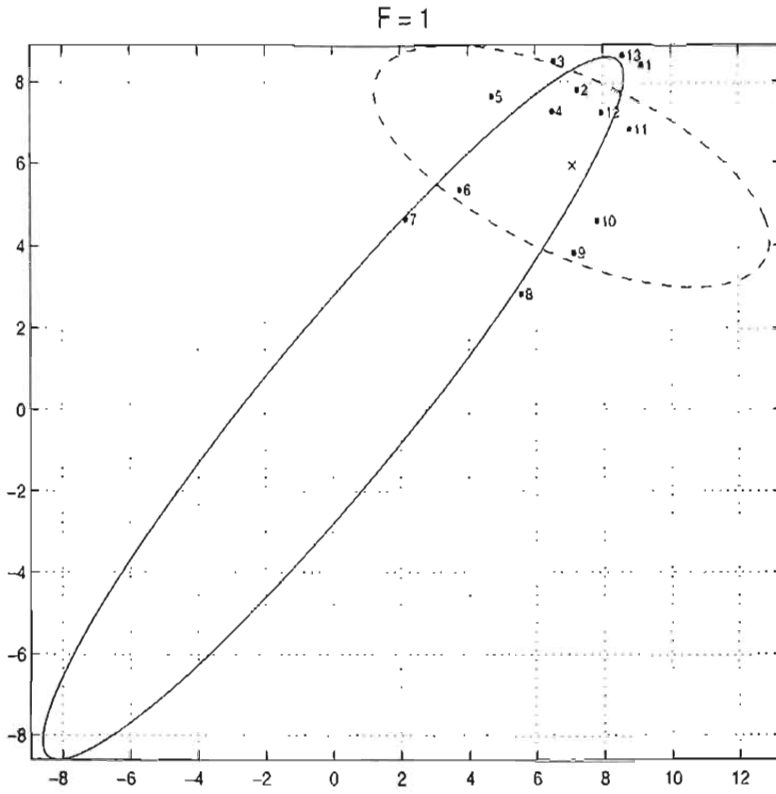
n°	x	y
1	8.485281	8.485281
2	8.035180	8.597102
3	7.264741	8.366182
4	6.204680	7.801730
5	4.897258	6.926247
6	3.394597	5.774637
7	1.756605	4.39281
8	5.224889	2.726102
9	6.478402	4.292694
10	7.473641	5.688149
11	8.170929	6.856836
12	8.542468	7.752162
13	8.573446	8.338434

Tabla 3.8: Puntos de prueba



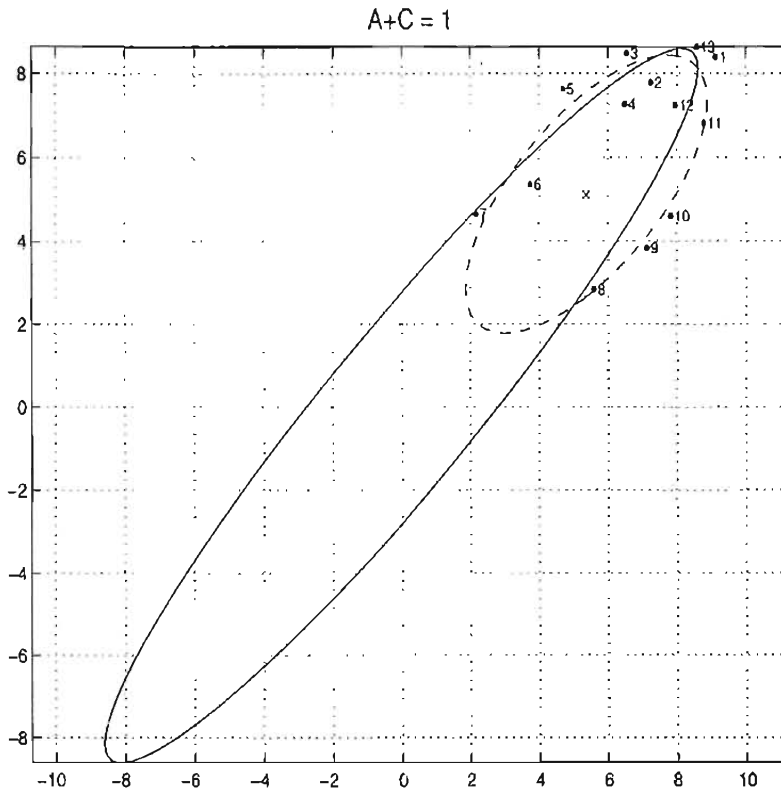
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	12	2	0.9860
Obtenida	(6.9894,5.9691)	5.7896	2.2046	0.9247

Figura 3.30: Resultados usando el método *Paton*



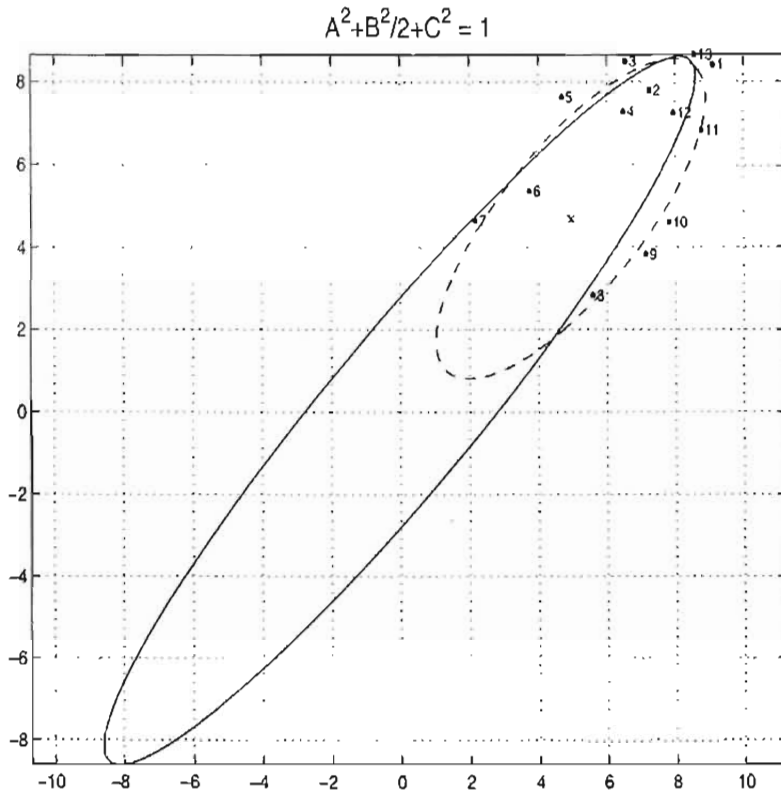
Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	12	2	0.9860
Obtenida	(7.0793,5.9415)	6.1575	2.2002	0.9340

Figura 3.31: Resultados usando el método *Rosin*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	12	2	0.9860
Obtenida	(5.3695,5.1159)	4.4284	1.9419	0.8987

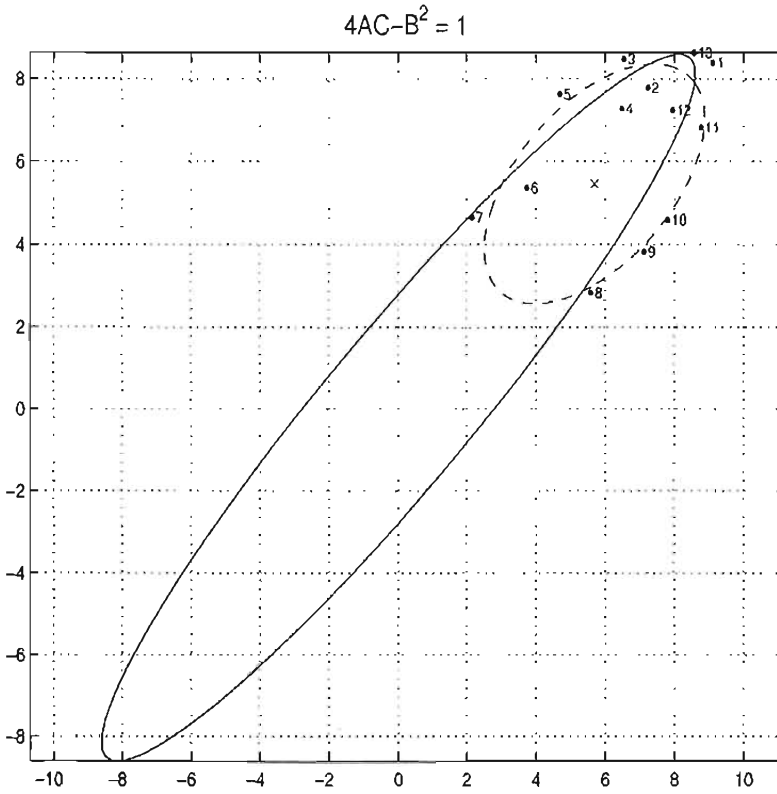
Figura 3.32: Resultados usando el método *Gander*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	12	2	0.9860
Obtenida	(4.9651,4.6795)	5.1711	1.9055	0.9296

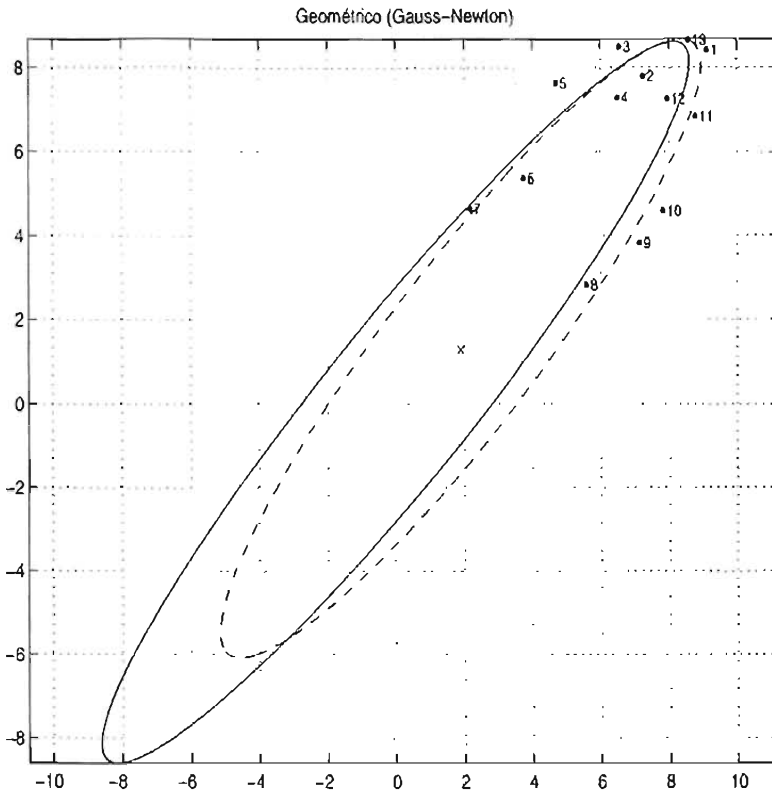
Figura 3.33: Resultados usando el método *Bookstein*

ESTA TESIS NO SALE
DE LA BIBLIOTECA



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	12	2	0.9860
Obtenida	(5.6963,5.4673)	3.7825	2.0482	0.8407

Figura 3.34: Resultados usando el método *Fitzgibbon*



Elipse	Centro	Semieje mayor	Semieje menor	Excentricidad
Referencia	(0,0)	12	2	0.9860
Obtenida	(1.8934,1.2954)	9.9914	2.0700	0.9783

Figura 3.35: Resultados usando el método *Geométrico*

Para los métodos *Paton* y *Rosin* se observa que las elipses cambiaron por completo, mientras que con *Gander*, *Bookstein* y *Fitzgibbon* se obtuvieron elipses menos alargadas. Para el método *Geométrico* la elipse obtenida es aceptable. Nuevamente se observó que los métodos basados en la distancia algebraica no resultan apropiados para el ajuste de elipses cuando los puntos pertenecen a las secciones de alta curvatura. En estos casos se obtienen mejores resultados aplicando un método basado en la distancia geométrica.

Conclusiones

En esta tesis se estudió el problema del ajuste de elipses. Se comenzó mostrando las diferentes formas en que se puede representar esta curva. De las representaciones, la ecuación general de segundo grado fue el modelo que se adoptó para llevar a cabo el ajuste por distancia algebraica y se pudo sacar ventaja de la forma en que se tienen los parámetros a estimar. Por otra parte, para el ajuste por distancia geométrica la representación que usamos fue la paramétrica ya que de esa manera es más sencillo minimizar esta distancia usando mínimos cuadrados no lineales.

El estudiar las cantidades que resultan invariantes bajo transformaciones euclidianas y las que no, sirvió para entender el comportamiento de la solución que se obtiene en el ajuste por distancia algebraica. De esta manera, es más sencillo ver porque normalizar con una cantidad que es invariante contribuye a que el ajuste también sea invariante. Por ejemplo, en el caso de la restricción que propone Fitzgibbon [9] el ajuste no sólo resulta invariante bajo transformaciones euclidianas, si no que también es específico de elipses y esto es una consecuencia directa del hecho de que se está forzando a que el discriminante resulte positivo.

En los experimentos del tercer capítulo se pudo comprobar como los métodos que no son invariantes encuentran elipses distintas cuando los puntos son transformados. También se pudo ver como el método propuesto por Fitzgibbon siempre encuentra una elipse a pesar de que los puntos no sigan una forma elíptica.

Otro comportamiento que se pudo observar con estos experimentos fue el sesgo por alta curvatura que presentan los métodos de ajuste por distancia algebraica. Se vio que con la distancia algebraica los puntos contribuyen de manera diferente a la estimación de los parámetros según sea su posición en la cónica. Por esta razón, la distancia algebraica deja de ser una buena aproximación de la distancia geométrica en las secciones de alta curvatura.

Con ayuda del sistema que se desarrolló en MATLAB como parte de este

trabajo, el lector puede llevar a cabo sus propios experimentos o bien repetir los que se han incluido para confirmar los resultados que se obtuvieron en este trabajo.

En la actualidad el problema del ajuste de elipses continúa siendo tema de investigación en diversas áreas y constantemente se publican artículos en revistas especializadas sobre éste. Aunque ya existen una gran cantidad de aplicaciones, surgen otras nuevas en áreas como la ortopedia o la medicina.

Este trabajo puede servir de introducción y referencia a aquellas personas que estén interesadas en el ajuste de elipses para aplicaciones como la graficación por computadora o el reconocimiento de patrones, o bien, para el ajuste de datos en otras áreas.

Cualquier comentario o duda acerca de este trabajo será muy bien recibido. Mi dirección de correo electrónico es: asd2ben@ciencias.unam.mx.

Apéndice A

Mínimos cuadrados

A.1 Introducción

El método de mínimos cuadrados resuelve el problema de ajustar sobre un conjunto de puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ en el plano, una forma funcional específica $y = f(x)$. La historia de los mínimos cuadrados se remonta a principios del siglo XIX cuando el gran matemático alemán Carl Friedrich Gauss (1777–1855) buscaba determinar la órbita del recién descubierto asteroide Ceres¹ (por equivocación, en un principio fue llamado planeta) a partir de las observaciones obtenidas por el astrónomo italiano Piazzi de Palermo en un período de 41 días.

Gauss resolvió teórica y numéricamente el problema de la determinación de las órbitas y desarrolló los fundamentos generales para la utilización de datos provenientes de muchas observaciones. Por medio de las ecuaciones obtenidas por él se pudo calcular con precisión la órbita de Ceres. Desde entonces la metodología utilizada por Gauss se conoce con el nombre de *método de mínimos cuadrados*.

La aplicación más clásica de este método es el ajuste polinomial (o regresión polinomial), que consiste en suponer que los datos guardan una relación del tipo

$$y_i = p(x_i) \quad p(x) \in \mathbb{P}^n$$

¹El asteroide Ceres fue descubierto por el astrónomo Piazzi de Palermo (1746–1826) en 1801 y bautizado por Johann Bode (1747–1826), director del observatorio de Berlín, quien lo identificó originalmente como un planeta. La confusión pudo deberse a que la regla de Titus-Bode predecía la existencia de una planeta a casi la misma distancia a la que se encuentra Ceres del Sol.

y el objetivo es encontrar los coeficientes de $p(x)$ que determinen la mejor aproximación. Aquí veremos que existen al menos dos maneras posibles de calcular esos parámetros, una apoyándonos del cálculo y la otra del álgebra lineal, además abordaremos el problema de los mínimos cuadrados lineales en general y mencionaremos algunas de las técnicas más utilizadas para resolver este problema. Comenzamos exponiendo las ideas básicas del ajuste lineal de datos.

A.2 Ajuste lineal por mínimos cuadrados

Supongamos que se tiene el siguiente conjunto de puntos

$$\begin{array}{c|c|c|c} x_1 & x_2 & \dots & x_n \\ \hline y_1 & y_2 & \dots & y_n \end{array} \quad (\text{A.1})$$

los cuales vistos sobre el plano aparecen así:



Figura A.1: Nube de puntos

Se busca ajustar una recta a estos puntos. Pensando de manera un poco ingenua, podríamos decir que existen muchas rectas dando el aspecto de ajustar bien a los datos pero ¿cuál de todas ellas es la mejor? No sería fácil decidirlo a simple vista. Es aquí donde surge la necesidad de contar con algún criterio que nos permita escoger a la mejor recta, la recta “ideal”.

Si pudiéramos hallar a y b tales que

$$y_i - (ax_i + b) = 0 \quad (\text{A.2})$$

para toda i no dudaríamos en que se ha encontrado la mejor recta. Por supuesto esto solo es posible si los puntos (x_i, y_i) están alineados. Al valor absoluto de (A.2) le llamaremos *residual* (r_i).

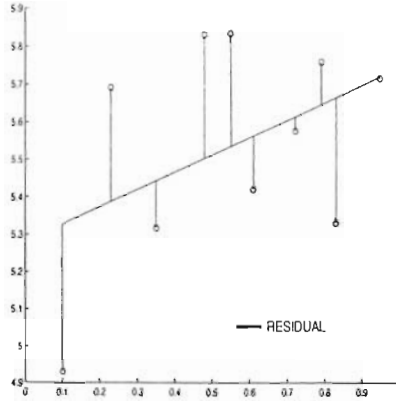


Figura A.2: Residuales

El problema de los mínimos cuadrados consiste en minimizar la suma de los cuadrados de los residuales. En el caso del ajuste lineal deseamos hallar el mínimo de la función

$$f(a, b) = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - ax_i - b)^2 \quad (\text{A.3})$$

donde r_i representa el i -ésimo residual. Ocupando la herramienta que el cálculo diferencial nos proporciona podemos hallar los parámetros a y b que minimizan f . Para ello derivamos (A.3) respecto de a y b , igualamos a cero y resolvemos el sistema de ecuaciones que resulta tal como se muestra a continuación.

$$\frac{\partial}{\partial a} \sum_{i=1}^n (y_i - ax_i - b)^2 = -2 \sum_{i=1}^n (y_i - ax_i - b)x_i = 0$$

$$\sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0 \quad (\text{A.4})$$

$$\frac{\partial}{\partial b} \sum_{i=1}^n (y_i - ax_i - b)^2 = -2 \sum_{i=1}^n (y_i - ax_i - b) = 0$$

$$\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - nb = 0 \quad (\text{A.5})$$

de (A.5) encontramos que

$$b = \bar{y} - a\bar{x} \quad (\text{A.6})$$

con lo que substituyendo (A.6) en (A.4) hallamos que

$$a = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \quad (\text{A.7})$$

Así quedan determinados los parámetros a y b a partir del conjunto de puntos satisfaciendo la condición de que la suma de los cuadrados de los residuales es mínima. Podemos afirmar lo anterior puesto que $\bar{x} = (a, b)$ es un punto crítico de la función f que cumple con

$$p \nabla^2 f(\bar{x}) p^t > 0 \quad p \in \mathbb{R}^2 \quad (\text{A.8})$$

es decir, la forma cuadrática en (A.8) es positiva definida por lo que satisface la condición de mínimo para extremos locales². Verifiquemos este hecho

²La notación $\nabla^2 f(x)$ representa la matriz simétrica de las segundas derivadas parciales de f en x a la cual se le da el nombre de Hessiano de f en x o matriz Hessiana de f en x . Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definimos el Hessiano de la función f en x como

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Sea $\bar{x} \in A$ un punto crítico de f entonces:

Si la forma cuadrática $x \nabla^2 f(\bar{x}) x'$ es positiva definida, f tiene un mínimo local en \bar{x} .

Si la forma cuadrática $x \nabla^2 f(\bar{x}) x'$ es negativa definida, f tiene un máximo local en \bar{x} .

realizando los cálculos correspondientes.

$$\begin{aligned}
 p\nabla^2 f(\bar{x})p^t &= (p_1 \ p_2) \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & 2n \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\
 &= \left(p_1 \sum_{i=0}^n x_i^2 + p_2 \sum_{i=0}^n x_i, \ p_1 \sum_{i=0}^n x_i + np_2 \right) \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\
 &= p_1^2 \sum_{i=0}^n x_i^2 + 2p_1 p_2 \sum_{i=0}^n x_i + np_2^2 \\
 &= \sum_{i=0}^n (p_1 x_i + p_2)^2 > 0
 \end{aligned}$$

\therefore la forma cuadrática en (A.8) es positiva definida.

De esta manera queda comprobado que $\bar{x} = (a, b)$ es el mínimo de la función f .

Ejemplo A.1 Ajuste lineal por mínimos cuadrados

Ahora aplicaremos los resultados obtenidos hasta el momento para llevar a cabo un ajuste lineal por mínimos cuadrados. Obtendremos los parámetros a y b de la recta a ajustar a partir de las fórmulas (A.6) y (A.7) que aparecen en las páginas anteriores. Supongamos que se cuenta con los siguientes datos:

x	4.6	3.4	3.8	4.2	4.9	5.1	5.3	5.8
y	5.6	4.9	5.3	5.1	6.1	5.8	6.3	6.2

Realizando los cálculos encontramos que:

$$a = 0.6049 \quad b = 2.8573$$

por lo que la recta de mínimos cuadrados es:

$$y = 2.8573 + 0.6049x$$

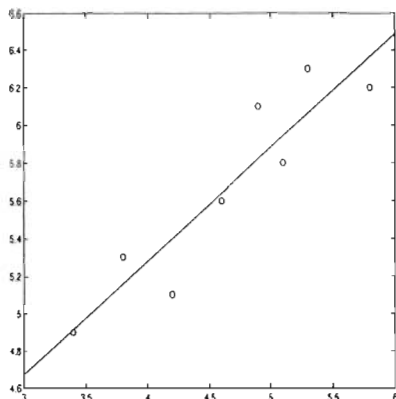


Figura A.3: Ajuste Lineal

Como se puede ver en este ejemplo una vez desarrollada la teoría resulta muy sencillo poner en práctica el ajuste lineal. A continuación veremos otra manera de minimizar los residuales, análoga a la que acabamos de ver, pero desde el enfoque del álgebra lineal.

A.3 Las ecuaciones normales

Supongamos nuevamente que se cuenta con un conjunto de puntos

$$\begin{array}{c|c|c|c} x_1 & x_2 & \dots & x_m \\ \hline y_1 & y_2 & \dots & y_m \end{array} \quad (\text{A.9})$$

y buscamos ajustar un polinomio de grado $n - 1$ por mínimos cuadrados a tal conjunto. Para cada una de las parejas (x_i, y_i) planteamos la ecuación $y_i = a_{n-1}x_i^{n-1} + a_{n-2}x_i^{n-2} + \dots + a_1x_i + a_0$. De las m parejas de puntos obtenemos un sistema de m ecuaciones y n incógnitas tal como el que aparece a continuación:

$$\left\{ \begin{array}{l} a_{n-1}x_1^{n-1} + a_{n-2}x_1^{n-2} + \dots + a_1x_1 + a_0 = y_1 \\ a_{n-1}x_2^{n-1} + a_{n-2}x_2^{n-2} + \dots + a_1x_2 + a_0 = y_2 \\ \vdots \\ a_{n-1}x_m^{n-1} + a_{n-2}x_m^{n-2} + \dots + a_1x_m + a_0 = y_m \end{array} \right\} \quad (\text{A.10})$$

donde a_{n-1}, \dots, a_0 son las incógnitas. Si el número de ecuaciones es mayor al de incógnitas se dice que el sistema está *sobredeterminado*. En general un sistema sobredeterminado *no tiene solución*.

Llamemos \mathbf{A} a la matriz de coeficientes y \mathbf{b} al vector columna de respuestas³. El sistema en (A.10) es equivalente a la ecuación matricial

$$\mathbf{Ax} = \mathbf{b} \tag{A.11}$$

$$\begin{bmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_m^{n-1} & x_m^{n-2} & \dots & x_m & 1 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

mientras que al vector columna de residuales lo llamaremos \mathbf{r} , en la notación matricial tenemos

$$\mathbf{r} = \mathbf{b} - \mathbf{Ax} \tag{A.12}$$

dado que, en general, el sistema (A.11) no tiene solución, debemos plantear el problema de manera diferente, esto es, buscamos

$$\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{Ax}\|_2^2 \tag{A.13}$$

de (A.13) conviene expresar

$$\begin{aligned} \|\mathbf{b} - \mathbf{Ax}\|_2^2 &= (\mathbf{b} - \mathbf{Ax})^t(\mathbf{b} - \mathbf{Ax}) \\ &= \mathbf{b}^t\mathbf{b} - 2\mathbf{x}^t\mathbf{A}^t\mathbf{b} + \mathbf{x}^t\mathbf{A}^t\mathbf{Ax} \end{aligned}$$

con cálculo diferencial podemos hallar el mínimo resolviendo

$$\begin{aligned} \frac{\partial}{\partial \mathbf{x}} (\mathbf{b}^t\mathbf{b} - 2\mathbf{x}^t\mathbf{A}^t\mathbf{b} + \mathbf{x}^t\mathbf{A}^t\mathbf{Ax}) &= 0 \\ -2\mathbf{A}^t\mathbf{b} + 2\mathbf{A}^t\mathbf{Ax} &= 0 \\ \mathbf{A}^t\mathbf{Ax} &= \mathbf{A}^t\mathbf{b} \end{aligned} \tag{A.14}$$

al sistema de ecuaciones en (A.14) se le conoce con el nombre de *ecuaciones normales*.

La matriz $\mathbf{B} = \mathbf{A}^t\mathbf{A}$ es simétrica y positiva definida. El método más común de resolver un sistema de ecuaciones de este tipo es usar la factorización Cholesky para $\mathbf{B} = \mathbf{R}^t\mathbf{R}$ donde \mathbf{R} es triangular superior. De esta manera (A.14) se convierte en:

$$\mathbf{R}^t\mathbf{Rx} = \mathbf{A}^t\mathbf{b}$$

³En la ecuación $ax_i + b = y_i$; a y y_i se le conoce como *respuesta*.

se resuelve

$$R^t w = A^t b \quad \text{para } w$$

y por último

$$Rx = w \quad \text{para } x$$

Ejemplo A.2 *Ajustando un polinomio de quinto grado por mínimos cuadrados usando la factorización de Cholesky*

Con ayuda de MATLAB hicimos una pequeña implementación. El código es el siguiente:

```
x=[-4.4009,...,3.9862]
y=[-4.2690,...,0.9942]
A=[x.^5 x.^4 x.^3 x.^2 x ones(length(x),1)]
B=A'*A
R=chol(B)
w=inv(R')*(A'*y)
parametros=inv(R)*w;
```

la variable `parametros` contiene los coeficientes del polinomio de quinto grado. En la tabla se muestran los puntos que se usaron para este ejemplo.

x	y	x	y
-4.4009	-4.2690	-3.8018	0.2339
-1.4055	1.6374	-0.6221	-2.8070
0.8986	3.5673	1.4747	0
2.7189	-4.9123	3.9862	0.9942

En la gráfica se ven el polinomio y los puntos

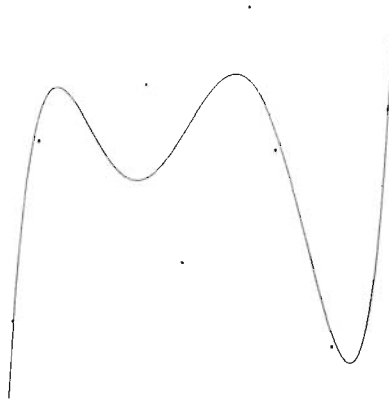


Figura A.4: Polinomio de quinto grado

Las ecuaciones normales siempre tienen solución y resuelven exactamente el problema de los mínimos cuadrados. Sin embargo, la principal desventaja al utilizarlas es que la condición de la matriz $\mathbf{A}^t\mathbf{A}$ es a menudo grande, siendo del orden del cuadrado de la condición de \mathbf{A} . Por este motivo las ecuaciones normales no se recomiendan como una buena aproximación al problema de mínimos cuadrados. Los métodos más confiables están basados en factorización de matrices usando matrices ortogonales. Estos son más caros, pero producen soluciones más exactas.

A.4 La factorización QR

La factorización QR permite obtener una solución al problema de mínimos cuadrados que será en la mayoría de los casos más exacta que la proporcionada por las ecuaciones normales. Al calcular la factorización QR reducida de \mathbf{A} obtenemos $\mathbf{A} = \mathbf{QR}$, donde \mathbf{Q} es una matriz *ortogonal*⁴ y \mathbf{R} es una matriz triangular superior. Entonces (A.11) se convierte en

$$\mathbf{QRx} = \mathbf{b}$$

multiplicando en ambos lados por \mathbf{Q}^t obtenemos

$$\mathbf{Rx} = \mathbf{Q}^t\mathbf{b}$$

se resuelve el sistema triangular superior

$$\mathbf{Rx} = \mathbf{y} \quad \text{donde} \quad \mathbf{y} = \mathbf{Q}^t\mathbf{b}$$

La ventaja de usar la descomposición QR es que preserva la norma de la solución por el hecho de que \mathbf{Q} es ortogonal, es decir:

$$\begin{aligned} \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{Ax}\|_2^2 &= \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{QRx}\|_2^2 \\ &= \min_{\mathbf{x}} \|\mathbf{QQ}^t\mathbf{b} - \mathbf{QRx}\|_2^2 \\ &= \min_{\mathbf{x}} \|\mathbf{Q}(\mathbf{Q}^t\mathbf{b} - \mathbf{Rx})\|_2^2 \\ &= \min_{\mathbf{x}} \|\mathbf{Q}\|_2^2 \|\mathbf{Q}^t\mathbf{b} - \mathbf{Rx}\|_2^2 \\ &= \min_{\mathbf{x}} \|\mathbf{Q}^t\mathbf{b} - \mathbf{Rx}\|_2^2 \end{aligned}$$

Ejemplo A.3 *Ajustando un polinomio de sexto grado por mínimos cuadrados mediante factorización QR*

El ejemplo que se muestra a continuación lo implementamos en un pequeño script de MATLAB; el código es el siguiente:

```
x=[0.6936; ...; 1.6871]
y=[-4.4483; ...; 1.1672]
A=[x.^6 x.^5 x.^4 x.^3 x.^2 x ones(length(x),1)]
[q,r]=qr(A,0)
parametros=inv(r)*(q'*y)
```

⁴Una matriz \mathbf{Q} de $n \times n$ se dice ortogonal si $\mathbf{Q}^t\mathbf{Q} = \mathbf{I}$, i.e. $\|\mathbf{Q}\|_2^2 = 1$

la variable `parametros` contiene los coeficientes del polinomio de sexto grado. Como se puede observar resulta relativamente sencillo implementar el ajuste polinomial; no es difícil ver de que manera hay que modificar la matriz `A` para ajustar polinomios de otros ordenes.

Los puntos para este ejemplo son:

x	y	x	y
0.6936	-4.4483	-15.7089	-5.3149
-18.0208	-1.5542	12.8609	-4.3188
-9.4079	0.9624	2.9191	5.5885
9.3372	0.2037	13.1647	8.8715
3.3268	-2.5286	5.1208	12.6913
0.5089	-10.2788	1.6871	1.1672

El polinomio se ve de la siguiente forma

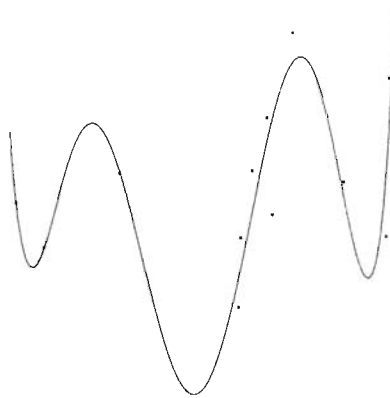


Figura A.5: Polinomio de sexto grado

Apéndice B

Mínimos cuadrados no lineales

B.1 Introducción

Un problema muy común es el de ajustar una función matemática a algunos datos experimentales variando los parámetros de la función. Sobre bases teóricas o empíricas se selecciona un modelo $F(\mathbf{x}, \mathbf{z})$ donde \mathbf{z} es el vector de las variables independientes y \mathbf{x} el vector de los n parámetros que serán ajustados hasta lograr que los datos sean explicados correctamente. Estas son las variables del problema.

Supongamos que se hacen m repeticiones de un experimento y que para cada \mathbf{z}_i se obtiene una correspondiente y_i . La mejor $F(\mathbf{x}, \mathbf{z})$ que ajuste a los datos será aquella en la que la diferencia entre los datos observados y los datos predecidos sea mínima.

$$\mathbf{r}_i(\mathbf{x}) = F(\mathbf{x}, \mathbf{z}_i) - y_i \quad i = 1, \dots, m$$

Como lo hicimos en el apéndice A, al valor absoluto de esta diferencia le llamaremos *residual*.

El problema de los mínimos cuadrados no lineales consiste en encontrar el mínimo global de la suma de los cuadrados de m funciones *no lineales*; es decir,

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{R}(\mathbf{x}) = \sum_{i=1}^m r_i^2(\mathbf{x}) = \| r(\mathbf{x}) \|_2^2$$

donde $r(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^m = [r_1(\mathbf{x}), \dots, r_m(\mathbf{x})]^t$ es el vector de residuales y cada $r_i(\mathbf{x}), i = 1, \dots, m$, es una función no lineal de \mathbb{R}^n en \mathbb{R} .

Esto nos conduce al sistema de ecuaciones no lineales

$$\begin{aligned} \mathbf{F}(\mathbf{x}, \mathbf{z}_1) &= y_1 \\ \mathbf{F}(\mathbf{x}, \mathbf{z}_2) &= y_2 \\ &\vdots \\ \mathbf{F}(\mathbf{x}, \mathbf{z}_m) &= y_m \end{aligned} \tag{B.1}$$

en el que generalmente m es mayor que n . Este sistema sobredeterminado no tiene una solución exacta pero es posible hallar una solución aproximada resolviendo el problema de los mínimos cuadrados no lineales.

B.2 Solución numérica del problema

El punto de partida para los algoritmos especializados en mínimos cuadrados no lineales es el método de Newton.

$$(\mathbf{J}_k^t \mathbf{J}_k + \mathbf{S}_k) \mathbf{p}_k = -\mathbf{J}_k^t \mathbf{f}_k \tag{B.2}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k \tag{B.3}$$

El principal problema de aplicar el método de Newton es el cálculo del término \mathbf{S}_k que involucra la evaluación de $mn(n+1)/2$ segundas derivadas. De aquí se desprenden dos clases de algoritmos: aquellos que ignoran el término \mathbf{S}_k que son llamados *algoritmos de residual pequeño* y aquellos que lo aproximan de alguna manera los cuales son llamados *algoritmos de residual grande*. Los métodos de Gauss-Newton y Levenberg-Marquardt corresponden a aquellos que ignoran el término \mathbf{S}_k , comenzaremos explicando estos métodos.

B.2.1 El método de Gauss-Newton

Ignorando el término \mathbf{S}_k en la ecuación (B.2) esta se convierte en

$$\mathbf{J}_k^t \mathbf{J}_k \mathbf{p}_k = -\mathbf{J}_k^t \mathbf{f}_k \tag{B.4}$$

que junto con

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k \tag{B.5}$$

definen el método de Gauss-Newton. Debido al mal condicionamiento de las ecuaciones normales en (B.4) se recurre a la factorización QR para resolver dicho sistema.

B.2.2 El método de Levenberg-Marquardt

El algoritmo de Levenberg-Marquardt es un algoritmo iterativo de optimización en el que el método de iteración presenta una ligera modificación sobre el método tradicional de Newton.

Este método evita las dificultades que el método de Gauss-Newton experimenta cuando a lo largo del proceso iterativo, en algún punto, la matriz Jacobiana no tiene rango completo además de ser un algoritmo efectivo para problemas de residual pequeño. La ecuación (B.2) se modifica para convertirse en

$$(\mathbf{J}_k^t \mathbf{J}_k + \mu_k \mathbf{I}) \mathbf{p}_k = -\mathbf{J}_k^t \mathbf{f}_k \quad (\text{B.6})$$

donde $\mu_k \geq 0$ es un escalar e \mathbf{I} es la matriz unitaria de orden n

Apéndice C

Sistema interactivo de ajuste de elipses

C.1 Introducción

El Sistema Interactivo de Ajuste de Elipses (SIAE) permite llevar a cabo el ajuste de elipses a un conjunto de puntos usando los métodos descritos en este trabajo. Está compuesto por un conjunto de `m-files` los cuales se ejecutan desde el archivo `principal.m`. Para comenzar a trabajar con el SIAE los archivos se deben alojar en una carpeta; por ejemplo, en la carpeta `C:\MATLABR11\work\siae`. Después, desde el shell de MATLAB tecleamos las siguientes instrucciones:

```
>> cd C:\MATLABR11\work\siae
>> principal
```

entonces, aparecerá una pantalla de presentación y enseguida se mostrará el menú principal, ver figura (C.1).

Elije alguna de las siguientes opciones:

- 1) **Generar los puntos para trabajar**
- 2) **Transformar los puntos**
- 3) **Ver los puntos generados**
- 4) **Guardar los puntos actuales**
- 5) **Guardar la gráfica actual**
- 6) **Ajustar una elipse a los puntos actuales**
- 7) **Ajustar una elipse de forma interactiva**
- 8) **Activar/Desactivar etiquetas [0]**
- 9) **Limpiar la pantalla**
- 10) **Salir**

Elijo la opción...

Figura C.1: Menú principal

C.2 Descripción del sistema

A continuación, se describen cada una de las opciones que aparecen en el menú principal.

- 1) **Generar los puntos para trabajar.** Despliega un submenú en el cual se especifican las distintas formas en las que se pueden generar los puntos. Este submenú será descrito en su totalidad más adelante.
- 2) **Trasformar los puntos.** Despliega un submenú con las distintas formas en que se pueden transformar los puntos. Será descrito más adelante.
- 3) **Ver los puntos actuales.** Al elegir esta opción se graficarán los puntos actuales. Los puntos deben ser generados primero.
- 4) **Guardar los puntos actuales.** Guarda los puntos actuales en un archivo con la extensión `.txt` dentro de la carpeta `puntos`. El formato

de escritura es el siguiente: en primer lugar se indica el número de puntos y posteriormente en dos columnas, una para las abcisas y otra para las ordenadas, se escriben las parejas de puntos. Ejemplo:

```
n
x_1  y_1
:    :
x_n  y_n
```

- 5) Guardar la gráfica actual. Guarda la figura actual en una imagen con formato .ps, .eps o .png.
- 6) Ajustar una elipse a los puntos actuales. Despliega un submenú donde se muestran los métodos disponibles para ajustar una elipse.
- 7) Ajustar una elipse en forma interactiva. Esta opción despliega una ventana en la cual se pueden ir agregando puntos y se obtiene una elipse por cada nuevo punto agregado a partir de cinco. Más adelante se verá un ejemplo de como usar esta opción.
- 8) Activar/Desactivar etiquetas. Al elegir esta opción los puntos serán etiquetados con un número lo que permitirá distinguir entre ellos.
- 9) Limpiar la pantalla. Borra todas las gráficas que estén sobre la figura actual.
- 10) Salir. Termina el programa.

Generar los puntos para trabajar

Al elegir la opción Generar los puntos para trabajar aparecerá el siguiente submenú:

Elije alguna de las siguientes opciones:

- 1) Leer los puntos de un archivo existente**
- 2) Generar los puntos con ayuda del mouse**
- 3) Generar puntos sobre una elipse**
- 4) Generar puntos sobre una elipse con aleatoriedad**
- 5) Volver al menú anterior**

Elijo la opción...

En él se muestran las opciones con que se cuenta para generar los puntos.

- 1) Leer los puntos de un archivo existente. Despliega una lista de archivos y carga al sistema los puntos contenidos en el archivo que se le indique en línea.
- 2) Generar los puntos con ayuda del mouse. Esta opción permite obtener los puntos con ayuda del mouse y de forma libre sobre la ventana que aparece; para finalizar éste proceso se presiona el botón derecho del mouse.
- 3) Generar los puntos sobre una elipse. Se genera una elipse usando la representación paramétrica. El usuario debe indicar las coordenadas del centro, la longitudes de los semiejes así como la inclinación de los mismos dada en radianes.
- 4) Generar los puntos sobre una elipse con aleatoriedad. Presenta las mismas características que la opción anterior, pero adiciona ruido gaussiano en cada una de las coordenadas suavizándolo por un factor entre 0 y 1.

Transformar los puntos

Contando con los puntos de trabajo en memoria, estos se pueden transformar mediante una traslación, una rotación o perturbándolos con ruido gaussiano según se indique en el siguiente submenú.

Elije alguna de las siguientes opciones:

- 1) Trasladar los puntos**
- 2) Rotar los puntos**
- 3) Agregar ruido gaussiano a los puntos**
- 4) Volver al menú anterior**

Elijo la opción...

- 1) **Trasladar los puntos.** Traslada los puntos según la cantidad de desplazamiento que se indique para cada coordenada.
- 2) **Rotar los puntos.** Rota los puntos según el ángulo en radianes que indique el usuario.
- 3) **Agregar ruido gaussiano a los puntos.** Agrega ruido gaussiano a cada una de las coordenadas de los puntos de trabajo suavizándolo con un factor entre 0 y 1.

Ajustar una elipse a los puntos actuales

Se puede ajustar una elipse a los puntos actuales aplicando cualquiera de los métodos que se estudian en este trabajo. En el siguiente submenú se elige el método deseado para llevar a cabo el ajuste.

Elige alguna de las siguientes opciones:

- 1) Distancia algebraica con restricción $\|a\|^2=1$ (Paton)
- 2) Distancia algebraica con restricción $F=1$ (Rosin)
- 3) Distancia algebraica con restricción $A+C=1$ (Gander)
- 4) Distancia algebraica con restricción $A^2+B^2/2+C^2=1$ (Bookstein)
- 5) Distancia algebraica con restricción $4AC-B^2=1$ (Fitzgibbon)
- 6) Distancia geométrica en forma paramétrica (Geométrico)
- 7) Volver al menú anterior

Elijo la opción...

C.3 Ejemplos

C.3.1 Comparando los métodos *Paton* y *Fitzgibbon*

A continuación se muestra un ejemplo de como trabajar con el sistema. En el menú principal se elige la opción Activar/Desactivar etiquetas y posteriormente la opción Generar los puntos para trabajar. En el siguiente

Elige alguna de las siguientes opciones:

- 1) Generar los puntos para trabajar
- 2) Transformar los puntos
- 3) Ver los puntos generados
- 4) Guardar los puntos actuales
- 5) Guardar la gráfica actual
- 6) Ajustar una elipse a los puntos actuales
- 7) Ajustar una elipse de forma interactiva
- 8) Activar/Desactivar etiquetas [1]
- 9) Limpiar la pantalla
- 10) Salir

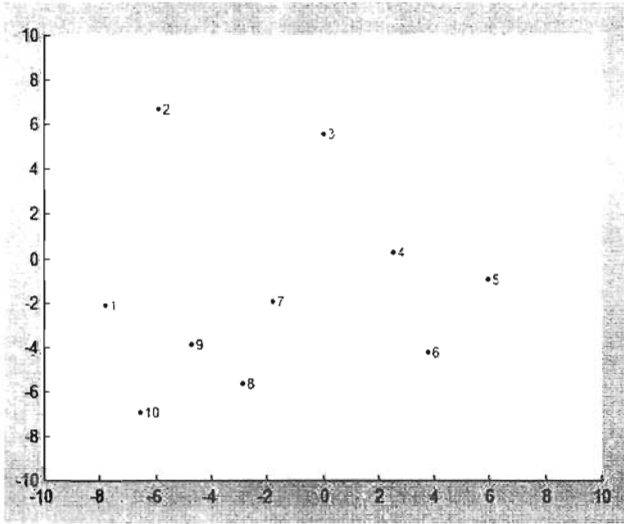
Elijo la opción... 1

menú se elige la opción Generar los puntos con ayuda del mouse y se toman diez puntos sobre la pantalla que aparece.

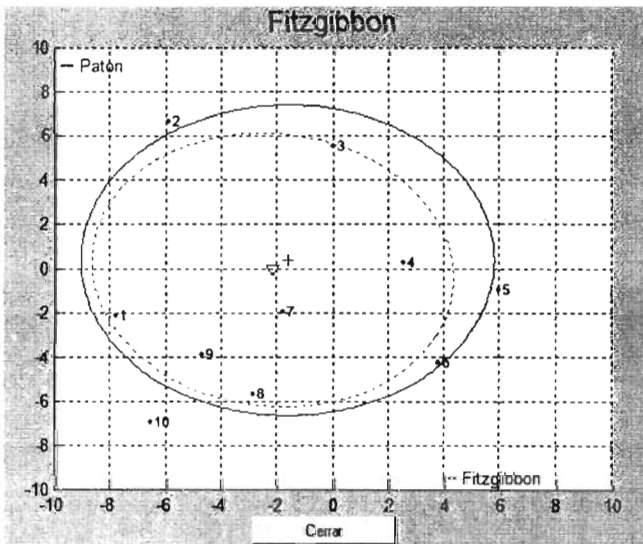
Elige alguna de las siguientes opciones:

- 1) Leer los puntos de un archivo existente
- 2) Generar los puntos con ayuda del mouse
- 3) Generar puntos sobre una elipse
- 4) Generar puntos sobre una elipse con aleatoriedad
- 5) Volver al menú anterior

Elijo la opción... 2



Los números que aparecen junto con los puntos son las etiquetas. A continuación, en el menú principal se elige la opción **Ajustar una elipse** a los puntos actuales y en el siguiente submenú se elige la opción correspondiente al método *Patón* y posteriormente la opción correspondiente al método *Fitzgibbon*; de esta forma ambos se graficarán en la misma pantalla.



Las características geométricas de las elipses obtenidas por cada uno de estos métodos son las siguientes:

```
* Resultados de aplicar el método Paton

- Centro = (-1.611730,0.383718)

- Longitud del semieje a = 7.432296

- Longitud del semieje b = 7.023132

- Excentricidad = 0.327221

- Rotación de la elipse alpha = 0.810617

Presiona ENTER para continuar...
```

Figura C.2: Resultados con el método *Paton*

```
* Resultados de aplicar el método Fitzgibbon

- Centro = (-2.140775,-0.056968)

- Longitud del semieje a = 6.592395

- Longitud del semieje b = 6.091742

- Excentricidad = 0.382257

- Rotación de la elipse alpha = 1.105016

Presiona ENTER para continuar...
```

Figura C.3: Resultados con el método *Fitzgibbon*

C.3.2 Ajuste en forma interactiva

Otra forma de ajustar elipses experimentando con los distintos métodos, es mediante la opción Ajustar una elipse en forma interactiva.

Elije alguna de las siguientes opciones:

- 1) Generar los puntos para trabajar
- 2) Transformar los puntos
- 3) Ver los puntos generados
- 4) Guardar los puntos actuales
- 5) Guardar la gráfica actual
- 6) Ajustar una elipse a los puntos actuales
- 7) Ajustar una elipse de forma interactiva
- 8) Activar/Desactivar etiquetas [0]
- 9) Limpiar la pantalla
- 10) Salir

Elijo la opción... 7

Al seleccionar ésta opción aparece la siguiente pantalla con algunas instrucciones para el usuario:

Lea detenidamente las siguientes:

INSTRUCCIONES DE USO

*** Capacidades disponibles ***

- 1) Puedes elegir tantos puntos como quieras apretando el botón izquierdo del mouse.
- 2) Puedes terminar en cualquier momento presionando el botón derecho del mouse.
- 3) Puedes alternar entre los métodos presionando la barra de espaciado.
- 4) Al terminar puedes trabajar normalmente con los puntos: guardarlos, transformarlos, etc.

*** Restricciones importantes ***

- 1) Se te piden al menos 5 puntos para empezar a trabajar.
- 2) En principio estás restringido a $[-10,10] \times [-10,10]$.

Presiona ENTER para continuar...

Después de haber señalado con el mouse 5 puntos sobre la pantalla se presenta una elipse que los ajusta. Se pulsa la barra espaciadora hasta que el título de la ventana indique que el método de ajuste que se está usando es el *Geométrico*. Después de señalar algunos puntos más, se observa la elipse de la figura (C.4).

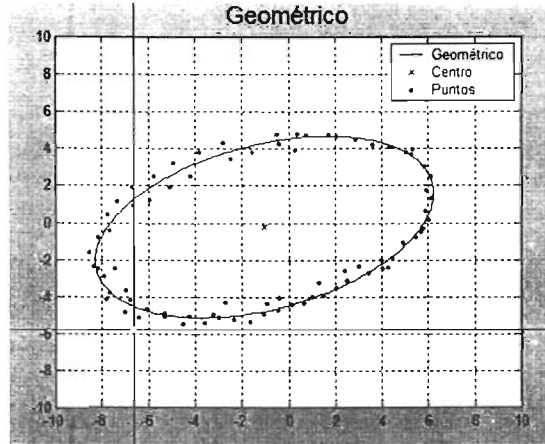


Figura C.4: Elipse que se obtiene con el método *Geométrico*

Una vez finalizado el proceso se muestran las características geométricas de la elipse así como el número de iteraciones necesario para obtener la solución:

▪ Resultados de aplicar el método Geométrico

El método convergió por criterio de residual pequeño

- Centro = (-1.634221, -0.213526)

- Longitud del semieje a = 7.687448

- Longitud del semieje b = 4.333310

- Excentricidad = 0.821912

- Rotación de la elipse alpha = -1.196391

- Nº de iteraciones = 6

C.4 Descripción de los m-files empleados

A continuación se describen los m-files de los módulos principales del sistema.

Módulo: Generar los puntos para trabajar

Este módulo cuenta con cuatro opciones:

- Leer los puntos de un archivo existente

Descripción: Lee de un archivo los puntos de trabajo.

Usa: leerPuntos.m

Entrada: Lista en pantalla los archivos .txt contenidos en el subdirectorio puntos y entonces solicita al usuario el nombre del archivo que contiene los puntos de prueba.

Salida: Al finalizar se obtiene en memoria los puntos de prueba.

- Generar los puntos con ayuda del mouse

Descripción: Captura en pantalla los puntos de trabajo.

Usa: obtenerPuntosMouse.m

Entrada: El área de trabajo $[x_{min}, x_{max}, y_{min}, y_{max}]$.

Salida: Al finalizar se obtiene en memoria los puntos de prueba.

- Generar puntos sobre una elipse

Descripción: Solicita al usuario las características geométricas de la elipse y genera los puntos de trabajo a partir de la representación paramétrica.

Usa: puntosElipse.m

Entrada: Las características geométricas de la elipse a representar: centro, longitud de los semiejes y rotación de la elipse en radianes.

Salida: Al finalizar se obtiene en memoria los puntos de prueba.

- Generar puntos sobre una elipse con aleatoriedad

Descripción: Solicita al usuario las características geométricas de la elipse y un factor en el intervalo $[0, 1]$ para suavizar el ruido; genera los puntos de trabajo a partir de la representación paramétrica.

Usa: puntosElipsePerturbada.m

Entrada: Número de puntos, centro, longitud de los semiejes, intervalo para el ángulo, factor para el ruido, rotación de la elipse en radianes.

Salida: Al finalizar se obtiene en memoria los puntos de prueba.

Módulo: Transformar los puntos

- Trasladar los puntos

Descripción: Traslada los puntos de trabajo según el desplazamiento que se indique para cada coordenada.

Usa: canalizarOpcMTP.m

Entrada: Solicita al usuario el desplazamiento para cada coordenada.

Salida: Al finalizar se obtiene en memoria los puntos de trabajo trasladados.

- Rotar los puntos

Descripción: Rota los puntos de trabajo según el ángulo en radianes que se indique.

Usa: canalizarOpcMTP.m

Entrada: Solicita al usuario el ángulo de rotación en radianes.

Salida: Al finalizar se obtiene en memoria los puntos de trabajo rotados.

- Agregar ruido gaussiano a los puntos

Descripción: Añade ruido gaussiano en cada una de las coordenadas de los puntos de trabajo.

Usa: canalizarOpcMTP.m

Entrada: Solicita al usuario que indique el factor de suavizamiento (entre 0 y 1).

Salida: Al finalizar se obtiene en memoria los puntos de trabajo perturbados.

Módulo: Ajustar una elipse a los puntos actuales

- Distancia algebraica con restricción $\|a\|^2 = 1$ (Paton)

Descripción: Ajusta una elipse a los puntos de trabajo con el método *Paton*.

Usa: paton.m

Entrada: Utiliza la variable global X que contiene a los puntos de trabajo.

Salida: Los coeficientes de la ecuación general de segundo grado de la elipse ajustada.

- Distancia algebraica con restricción $F = 1$ (Rosin)

Descripción: Ajusta una elipse a los puntos de trabajo con el método *Rosin*.

Usa: rosin.m

Entrada: Utiliza la variable global X que contiene a los puntos de trabajo.

Salida: Los coeficientes de la ecuación general de segundo grado de la elipse ajustada.

- Distancia algebraica con restricción $A + C = 1$ (Gander)

Descripción: Ajusta una elipse a los puntos de trabajo con el método *Gander*.

Usa: gander.m

Entrada: Utiliza la variable global X que contiene a los puntos de trabajo.

Salida: Los coeficientes de la ecuación general de segundo grado de la elipse ajustada.

- Distancia algebraica con restricción $A^2 + B^2 + C^2 = 1$
(Bookstein)

Descripción: Ajusta una elipse a los puntos de trabajo con el método *Bookstein*.

Usa: bookstein.m

Entrada: Utiliza la variable global X que contiene a los puntos de trabajo.

Salida: los coeficientes de la ecuación general de segundo grado de la elipse ajustada.

- Distancia algebraica con restricción $4AC - B^2 = 1$
(Fitzgibbon)

Descripción: Ajusta una elipse a los puntos de trabajo con el método *Fitzgibbon*.

Usa: fitzgibbon.m

Entrada: Utiliza la variable global X que contiene a los puntos de trabajo.

Salida: Los coeficientes de la ecuación general de segundo grado de la elipse ajustada.

- Distancia geométrica en forma paramétrica (Geométrico)

Descripción: Ajusta una elipse a los puntos de trabajo con el método *Geométrico*; el usuario debe indicar con cual de los métodos algebraicos se va a calcular la aproximación inicial.

Usa: geometrico.m

Entrada: Utiliza la variable global X que contiene a los puntos de trabajo.

Salida: El centro, la longitud de los semiejes y la rotación de la elipse ajustada.

C.5 Distribución del código

El SIAE fue programado en MATLAB 5.3 (versión 5.3.0.10183 (R11)) y ha sido probado exitosamente en las versiones 6.5 y 7.0 de MATLAB.

Los `m-files` que conforman el sistema junto con los archivos de texto plano que contienen los puntos de prueba que se utilizaron en los experimentos del Capítulo 3, vienen dentro del CD que se incluye junto con la tesis, o bien, el lector puede solicitarlos directamente al autor en la dirección de correo electrónico: `asd2ben@ciencias.unam.mx`.

Bibliografía

- [1] Albano, A. *Representation of digitized contours in terms of conic arcs and straight-line segments*. Computer Graphics and Image Processing 3, 23-33, 1974.
- [2] Ballard, D.H. y Brown C.M. *Computer Vision*. Prentice-Hall, Englewood Cliffs, N.J. 1982.
- [3] Biggerstaff, R.H. *Three variations in dental arch form estimated by a quadratic equation*. Journal of Dental Research 51, 1509, 1972.
- [4] Bookstein, F.L. *Fitting conic sections to scattered data*. Computer Vision, Graphics and Image Processing 9, 56-71, 1979.
- [5] Brannan, D.A., Esplen, M.F. y Gray J.J. *Geometry* The Open University. Cambridge University Press, 1999.
- [6] Cooper, D.R. y Yalabik, N. *On the computational cost of approximating and recognizing noise-perturbed straight lines and quadratic arcs in the plane*. IEEE Trans. Computers C-25, 1976, 1020.
- [7] Dennis, J.E., Schnabel, R.B. *Numerical methods for unconstrained optimization and nonlinear equations*. Prentice-Hall Series in Computational Mathematics, 1983.
- [8] Fitzgibbon, A.W. y Fisher, R.B. *A buyer's guide to conic fitting*. In Proceedings of the British Machine Vision Conference, pages 265-271, Birmingham, 1995.
- [9] Fitzgibbon, A.W., Pilu M. y Fischer, R.B. *Direct least squares fitting of ellipses*. Technical Report DAIRP-794, Department of Artificial Intelligence, The University of Edinburgh, January 1996.

-
- [10] Fitzgibbon, A.W. *Stable Segmentation of 2D curves*. Ph. D. thesis, Dept. of Artificial Intelligence, University of Edinburgh, 1998.
- [11] Forsyth, D., Mundy, J.L., Zisserman, A., and Brown, C.M. *Projectively invariant representations using implicit algebraic curves*. Image and vision computing, 9(2):130–136, Abril 1991.
- [12] Forsyth, D., Mundy, J.L., Zisserman, A., et al. *Invariant descriptors for 3D object recognition and pose*. IEEE Transactions on Pattern Analysis and Machine Analysis, 13(10):971–991, 1991.
- [13] Gander, W. y von Matt, U. *Some least squares problems, solving problems in scientific computing using Maple and Matlab*. W. Gander and J.H. ed., 251-266, Springer-Verlag, 1993.
- [14] Gander, W., Golub, G.H. y Strebel R. *Least-squares fitting of circles and ellipses*. BIT, no. 43, pp. 558–578, 1994.
- [15] Gnanadesikan, R. *Methods for statistical data analysis of multivariate observations*. Wiley, New York, 1977.
- [16] Golub, G.H. y Van Loan, C.F. *Matrix Computations*. The Johns Hopkins University Press. 1989.
- [17] Halíř, R. y Flusser, J. *Numerically Stable Direct Least Squares Fitting Of Ellipses* En V. Škala, editor, Proc. of the 6th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media (WSCG'98), volumen 1, pags. 125–132, Plzeň República Checa, febrero 1998.
- [18] Kanatani, K. *Statistical bias of conic fitting and renormalization*. IEEE Transactions on Pattern Analysis and Machine Analysis, 16(3):320–326, 1994.
- [19] Kapur, D. y Mundy, J.L. *Fitting affine-invariant conics to curves*. En J.L. Mundy y A. Zisserman, editores, Geometric Invariance in Computer Vision. MIT Press, 1992.
- [20] Kerrigan*, S., McKenna* S.J., Ricketts* I.W. y Wigderowitz** C. *Eccentric elliptical contours in total hip replacements* *Applied Computing, University of Dundee, DD1 4HN. **Orthopaedics and Trauma Surgery, Ninewells Hospital, Dundee, DD1 9SY. Presentado en Medical Image Understanding and Analysis 2004.

-
- [21] Lehmann, Charles H. *Geometría Analítica*. Editorial Limusa. Decimocuarta reimpresión. 1990.
- [22] The Mathworks Inc. MATLAB: sistema para cálculo numérico y visualización. <http://www.mathworks.com>
- [23] Pajares. M.G., de la Cruz. G.J. *Visión por computador: imágenes digitales y aplicaciones*. RA-MA Editorial; Alfaomega Grupo Editor. 2002.
- [24] Paton, K.A. *Conic sections in automatic chromosome analysis*. En B. Meltzer y D. Michie, editores, Machine Intelligence 5, 411-434. Edinburgh University Press, Edinburgh, 1969.
- [25] Ramírez, Ana I. *Geometría analítica: una introducción a la geometría*. Coordinación de Servicios Editoriales, Facultad de Ciencias, UNAM. 1998.
- [26] McLeod, Robin J.Y. y Baart, M. Louisa *Geometry and Interpolation of curves and surfaces*. Cambridge University Press, 1998.
- [27] Rosin, P.L. *A note on the least squares fitting of ellipses*. Pattern Recognition Letters, no. 14, pp. 799-808, Oct. 1993.
- [28] Sampson, P.D. *Fitting conic sections to "very scattered" data: An iterative refinement of the Bookstein algorithm*. Computer Graphics and Image Processing, 18:97-108, 1982.
- [29] Taubin, G. *Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation*. IEEE Transactions on Pattern Analysis and Machine Analysis, 13(11):1115-1138, noviembre 1991.
- [30] Wooton W., Beckenbach E.F., Fleming F.J. *Modern Analytic Geometry*. Houghton Mifflin Company, 1972.
- [31] Zhang, Z. *Parameter estimation techniques: a tutorial with application to conic fitting*. Institut National de Recherche en Informatique et en Automatique. Rapport de recherche n° 2676. Octobre 1995, 44 pages.