

Clase 27. Matriz Positiva Definida.

Curso ‘Linear Algebra’ del MIT.

Resumen

En el mundo de las matrices simétricas existe un tipo especial llamadas Matrices Positivas Definidas. Sus características, sumadas a las de las simétricas, se vinculan con varios temas que hemos estudiado como los eigenvalores, pivotes y determinantes. A esto se suma un valor adicional que nos permite evaluarlas directamente. En esta ocasión revisaremos en detalle dichas cualidades de ellas.

1. Matrices positivas definidas.

Si las matrices simétricas tienen muy buenas cualidades para trabajar con ellas, existe una subclase de éstas que son aún mejor, las cuales reciben el nombre de **Matrices Positivas Definidas** y se caracterizan por lo siguiente:

1. Ser simétricas.
2. Todos sus eigenvalores son **positivos**.
3. Todos sus pivotes son **positivos**.
4. Todos sus subdeterminantes son **positivos**.

Los puntos **1**, **2** y **3** se vinculan con lo que estudiamos en la Clase 25. Si una matriz es **simétrica**, podemos asegurar que la cantidad de entradas pivotes positivas y eigenvalores positivos serán iguales en ella. Para las matrices positivas definidas sucede que todas las pivotes y, como consecuencia, todos los eigenvalores, son mayores a cero.

En cuanto al punto **4**, podría ser una opción considerar el determinante de toda una matriz simétrica y que esta fuese positiva. Recordemos que este escalar es el producto de las entradas pivotes y también la multiplicación entre los eigenvalores. El problema es que dichos valores

pueden ser menores a cero y el resultado de aquella operación seguir siendo positivo. En ese sentido, como alternativa se calculan **determinantes de submatrices** ubicadas en la **parte superior izquierda** de la inicial y, en las matrices positivas definidas, todos son positivos.

Para entenderlo mejor, veamos la siguiente matriz positiva definida:

$$A = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix}$$

Como se puede apreciar, es simétrica.

$$A^T = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix} = A$$

En cuanto a sus entradas pivotes, son todas positivas.

$$A = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 3 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

Por lo tanto, podemos asegurar que sus eigenvalores son positivos. En particular, corresponden a los siguientes¹:

$$\lambda_1 \approx 10.95 \quad \lambda_2 \approx 6.54 \quad \lambda_3 \approx 1.51$$

Con respecto a los determinantes de las submatrices superiores izquierdas, nos referimos a lo siguiente: Primero tomamos la que está en la esquina de dicha ubicación. La denominaremos como A_1 .

$$A_1 = \begin{bmatrix} 4 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \Rightarrow \det(A_1) = 4$$

Luego, calculamos el determinante de la matriz A_2 que será una de 2×2 con las entradas superiores izquierdas de A .

$$A_2 = \begin{bmatrix} 4 & 2 & \cdot \\ 2 & 10 & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \Rightarrow \det(A_2) = \begin{vmatrix} 4 & 2 \\ 2 & 10 \end{vmatrix} = 36$$

¹Los calculé con el programa R mediante la función `eigen`.

Finalmente, la submatriz A_3 será toda la matriz A , cuyo determinante lo podemos calcular como el producto de las entradas pivotes de esta última.

$$A_3 = A = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix} \implies \det(A_3) = \det(A) = 4 \cdot 9 \cdot 3 = 108$$

Como vemos, tanto los determinantes de las submatrices como el de la matriz A son positivos, pero puede darse el caso donde esta igualdad de signos no ocurra. Por ejemplo, en la siguiente matriz:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}$$

su determinante es:

$$\det(A) = (-1) \cdot (-3) = 3 > 0$$

No obstante, el signo del determinante de uno de sus submatrices es negativo.

$$A_1 = [-1] \implies \det(A_1) = -1 < 0$$

Es más, sus pivotes y eigenvalores son menores a cero, donde en ambos casos son -1 y -3 . Por ello es que se consideran a los subdeterminantes de la matriz A para evaluar si la matriz es positiva definida o no.

2. Evaluando si una matriz es positiva definida.

En la sección anterior vimos algunas condiciones para evaluar si una matriz simétrica es positiva definida o no. Formalicemos aquello.

Sea A la siguiente matriz simétrica de 2×2 .

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

En primer lugar, será positiva definida si sus **eigenvalores**:

$$\lambda_1, \lambda_2 > 0$$

También, si sus **subdeterminantes** $\det(A_1) = a$ y $\det(A_2) = ac - b^2$ (este último siendo

igual al $\det(A)$):

$$a, ac - b^2 > 0$$

En cuanto a los pivotes de A , sabemos que el primero es $a_{1,1} = a$. El segundo $a_{2,2}$ podemos obtenerlo multiplicando la primera fila por $(-b/a)$ para, posteriormente, sumar la segunda a ésta.

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \Rightarrow \begin{bmatrix} a & b \\ 0 & c - \frac{b^2}{a} \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & \frac{ac - b^2}{a} \end{bmatrix}$$

Por lo tanto, para que A sea positiva definida, debe cumplirse que sus **entradas pivotes**:

$$a, \frac{ac - b^2}{a} > 0$$

Veamos que esto último es equivalente a:

$$\det(A_1), \frac{\det(A_2)}{\det(A_1)} > 0$$

Y el último pivote para una matriz de $n \times n$ se puede replicar como $\frac{\det(A_n)}{\det(A_{n-1})}$.

En síntesis, una matriz $A = A^T$ es positiva definida si:

1. Eigenvalores $\Rightarrow \lambda_1, \lambda_2 > 0$.
2. Subdeterminantes $\Rightarrow a, ac - b^2 > 0$.
3. Pivotes $\Rightarrow a, \frac{ac - b^2}{a} > 0$.

2.1 Forma cuadrática.

No obstante, la **prueba que es comúnmente usada** para evaluar si una matriz simétrica es **positiva definida** o no, está dada a partir de la siguiente función vectorial:

$$f(\vec{x}) = \vec{x}^T A \vec{x}; \quad \forall \vec{x} \neq \vec{0}$$

Esta función recibe el nombre de **Forma Cuadrática** puesto que, algebraicamente, corresponde a un polinomio de grado dos². Por ejemplo, sean:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

²De hecho, recordemos que $A\vec{x}$ nos entrega un polinomio lineal.

Entonces:

$$\vec{x}^T A \vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

Se señala que una matriz **simétrica** $A = A^T$ es **positiva definida** si:

$$\vec{x}^T A \vec{x} > 0; \quad \forall \vec{x} \neq \vec{0}$$

3. Matriz positiva semidefinida.

No siempre las matrices simétricas serán positivas definidas. En particular, aquellas donde al menos uno de sus eigenvalores es cero y su forma cuadrática es no negativa, se catalogan como **Positivas Semidefinidas**. Es decir, si:

1. La matriz A es Simétrica (i.e, $A = A^T$).
2. Eigenvalores de A son $\lambda_i \geq 0$.
3. Al menos una de sus columnas no tiene una entrada pivote.
4. $\vec{x}^T A \vec{x} \geq 0$.

Entonces decimos que A es positiva semidefinida.

Algo no menor para evaluarlas es que uno de sus eigenvalores sea $\lambda_i = 0$, puesto que esto ocurre cuando las matrices son **singulares**. En otras palabras, cuando $\det(A) = 0$. Pero el resto de dichos escalares deben ser **positivos** para que la matriz sea positiva semidefinida.

4. Matrices positivas definidas y el mínimo de la forma cuadrática.

Una aplicación de la forma cuadrática $\vec{x}^T A \vec{x}$ y las matrices positivas definidas es al resolver problemas de optimización. En este caso, para encontrar el **valor mínimo** de una función de varias variables.

Un concepto que agregaremos antes, es el de las **matrices negativas definidas** que son aquellas donde $\vec{x}^T A \vec{x} < 0$ para todo los $\vec{x} \neq \vec{0}$. Cumplen las mismas características que las positivas definidas (eigenvalores, pivotes, subdeterminantes), pero en signo negativo.

Digamos que $f(\vec{x})$ es una forma cuadrática. En general, se cumple que si:

1. $f(\vec{x}) > 0$ entonces la forma cuadrática es **positiva definida** para todo $\vec{x} \neq \vec{0}$.
2. $f(\vec{x}) < 0$ entonces la forma cuadrática es **negativa definida** para todo $\vec{x} \neq \vec{0}$.
3. $f(\vec{x}) > 0$ y $f(\vec{x}) < 0$ entonces la forma cuadrática es **indefinida**.

Lo anterior quiere decir que, por ejemplo, cuando la forma cuadrática y, por consiguiente, la matriz simétrica A es **positiva definida**, la gráfica de la primera es **creciente** para todos los componentes de $\vec{x} \neq \vec{0}$. Por lo tanto, el vector $\vec{x} = \vec{0}$ corresponde al **valor mínimo** de la función.

Para ser más preciso, gráficamente las **matrices positivas definidas** nos permiten obtener funciones cuya superficie es **estrictamente convexa** donde el **valor mínimo** se ubica en su centro, como lo vemos a continuación para $\vec{x} \in \mathbb{R}^2$, $A \in \mathbb{R}^{2 \times 2}$ y donde $z = f(x, y)$.

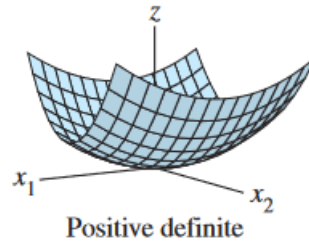


Figura 1: Lay (2010). *Linear Algebra and Its Applications*. Pp. 406.

En cuanto a las **matrices negativas definidas** y a aquellas **indefinidas**, el centro de las primeras será su **valor máximo**, mientras que en las segundas aquel punto será uno de **ensillado** (*saddle point*), es decir, para algunas variables será una valor mínimo y para otras uno máximo.

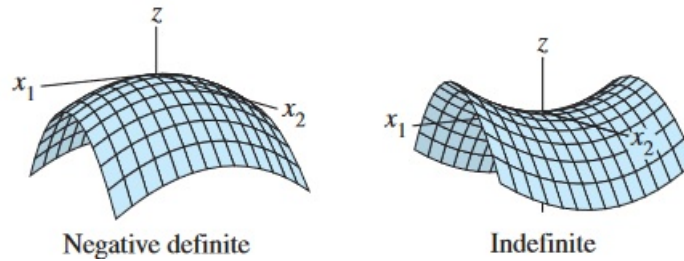


Figura 2: Lay (2010). *Linear Algebra and Its Applications*. Pp. 406.

Si, por ejemplo, necesitamos verificar que el centro de esta función estrictamente convexa sea su valor mínimo, evaluamos que sus **primeras derivadas** sean iguales a cero en dicho lugar y, fundamentalmente, que sus **segundas derivadas** sean mayores a cero para componentes de \vec{x} mayores de donde está su centro.

Es más, si la **Matriz Hessiana**³ de una función $f(\vec{x})$ es **positiva definida**, podemos tener por garantizado que tiene un **valor mínimo**. Para el caso de la **forma cuadrática** $f(\vec{x}) = \vec{x}^T A \vec{x}$, corresponderá a su centro donde la pendiente del plano tangente en dicho lugar es cero.

Esto nos muestra el poder de saber si una matriz simétrica es positiva definida o no, para resolver problemas de optimización con solo ver cumple con las condiciones para serlo.

5. Relación entre la forma cuadrática y el método de eliminación de Gauss.

Para evaluar si una matriz simétrica es positiva definida, hemos usado dos métodos: el **método de eliminación de Gauss** y la **forma cuadrática**. Algo interesante a observar, es que ambos tienen un vínculo.

Digamos que:

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{y} \quad \vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Entonces:

$$\vec{x}^T A \vec{x} = ax^2 + bxy + cy^2$$

La expresión $ax^2 + bxy$ podemos usarla para **completar un cuadrado de binomio** de la siguiente manera:

$$a \left(x + \frac{b}{a}y \right)^2$$

Si lo expandimos, obtenemos que:

$$a \left(x + \frac{b}{a}y \right)^2 = a \left(x^2 + 2\frac{b}{a}xy + \frac{b^2}{a^2}y^2 \right) = a^2 + 2bxy + \frac{b^2}{a}y^2$$

Necesitamos encontrar un valor que sea igual a cy^2 para que el polinomio de arriba sea $\vec{x}^T A \vec{x}$. Aquello corresponde a la suma de $\frac{b^2}{a}y^2$ con otra expresión, que denotaremos como w y que

³La matriz de segundas derivadas de una función de varias variables.

debe ser igual a cy^2 .

$$\begin{aligned}\frac{b^2}{a}y^2 + w &= cy^2 \\ w &= cy^2 - \frac{b^2}{a}y^2 \\ w &= \left(\frac{ac - b^2}{a}\right)y^2\end{aligned}$$

Por lo tanto, la **forma cuadrática** para una matriz simétrica de 2×2 podemos expresarla también como:

$$\vec{x}^T A \vec{x} = a \left(x + \frac{b}{a}y\right)^2 + \left(\frac{ac - b^2}{a}\right)y^2$$

Esta forma de expresar la **forma cuadrática** muestra su **relación** con el **método de eliminación de Gauss**. El valor a que está afuera del binomio cuadrático y la constante $\left(\frac{ac - b^2}{a}\right)$ son los **pivotes** de la matriz A que se encuentran con este método. Por otra parte, $\left(\frac{b}{a}\right)$ es el múltiplo que usamos para pasar de $A \rightarrow U$ y que aparece en la matriz L de la **descomposición** $A = LU$, que es la expresión matricial de este algoritmo.

$$\begin{aligned}A &= LU \\ \begin{bmatrix} a & b \\ b & c \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ \frac{b}{a} & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ 0 & \frac{ac - b^2}{a} \end{bmatrix}\end{aligned}$$

Por lo tanto, al usar la forma cuadrática y al completar el cuadrado como lo acabamos de hacer, podemos encontrar los pivotes de la matriz simétrica y evaluar si es positiva definida según el signo de éstos y, por consiguiente, el de los subdeterminantes y eigenvalores.

6. Elipse y el Teorema del Eje Principal.

Otra característica de las matrices positivas definidas de 2×2 , es que la **curva de nivel** del paraboloide de su forma cuadrática en $f(\vec{x}) = k$ ($k = \text{constante}$), es una **elipse**. Es decir, si

$$\vec{x}^T A \vec{x} = ax^2 + 2bxy + cy^2 > 0,$$

entonces

$$\vec{x}^T A \vec{x} = ax^2 + 2bxy + cy^2 = k$$

es una elipse⁴.

⁴En particular, es una sección cónica.

Ahora bien, la ecuación $\vec{x}^T A \vec{x}$ no es una elipse estándar. La ecuación canónica de esta figura corresponde a:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

la cual es de ancho $2a$ y su altura es $2b$.

En los lados de la elipse puede darse que $a \geq b > 0$ o que $a \leq b > 0$, de manera que podemos encontrarnos con ella de dos formas como vemos en la siguiente imagen (sea $\alpha = a$ y $\beta = b$):

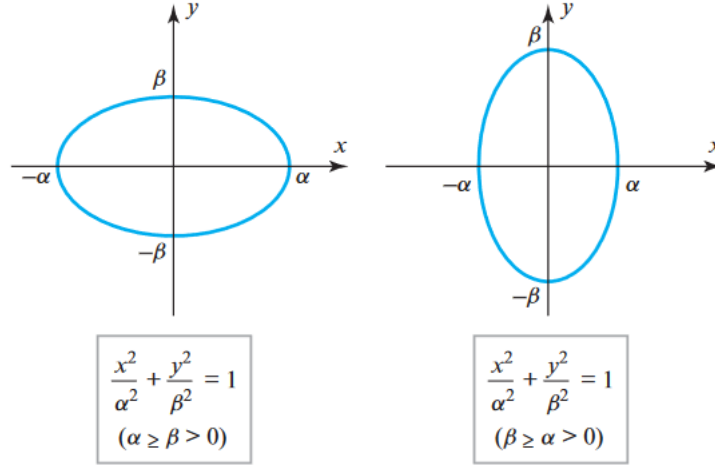


Figura 3: Anton y Rorres (2013). *Elementary Linear Algebra. Applications Version*. Pp 421.

En otras palabras, cuando obtenemos la forma canónica de la ecuación de una elipse, sus ejes mayor y menor⁵ coinciden con estar en los de x e y . Esto no ocurre en $\vec{x}^T A \vec{x}$, puesto que debido al término $2bxy$ (conocido como el **término de producto cruzado**), aquella gráfica se encontrará **inclinada**.

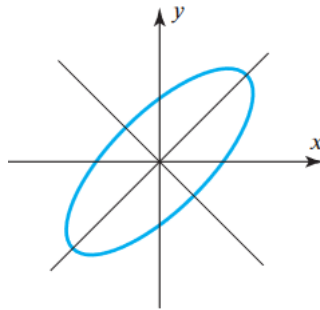


Figura 4: Anton y Rorres (2013). *Elementary Linear Algebra. Applications Version*. Pp 422.

Por lo tanto, al tener el término de producto cruzado, se hace más difícil trabajar con una función como $\vec{x}^T A \vec{x}$, tanto al graficarla como al modelar con ella. Para resolver esta

⁵Los segmentos perpendiculares que están en su interior, donde el “mayor” es el de más diámetro y el “menor”, el de menos.

problemática, podemos realizar un **cambio de ejes** que nos permita obtener los de la elipse inclinada y, para ello, debemos eliminar a $2bxy$.

Si aplicamos un poco de álgebra matricial, veremos que necesitamos que A sea **diagonal** para que el término $2bxy$ no aparezca en la forma cuadrática $\vec{x}^T A \vec{x}$. Por ejemplo:

$$\vec{x}^T A \vec{x} = \begin{bmatrix} x & y \end{bmatrix} \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + by^2 = k$$

En principio, la matriz A no es diagonal en la forma cuadrática $\vec{x}^T A \vec{x}$, pero tenemos la ventaja en que es **positiva definida** y, por consiguiente, **simétrica**. Por lo tanto, podemos aplicar el **Teorema Espectral** para descomponerla como:

$$A = Q \Lambda Q^T$$

donde Q es la matriz de eigenvectores columna ortonormalizados y Λ la matriz diagonal de eigenvalores, ambos provenientes de la matriz A .

Reemplacemos a $A = Q \Lambda Q^T$ en $\vec{x}^T A \vec{x}$.

$$\vec{x}^T A \vec{x} = \vec{x}^T (Q \Lambda Q^T) \vec{x}$$

Luego, para hacer el cambio de ejes sustituyamos $\vec{x} = Q \vec{y}$ en la ecuación de arriba.

$$\vec{x}^T A \vec{x} = (Q \vec{y})^T Q \Lambda Q^T (Q \vec{y}) = \vec{y}^T Q^T Q \Lambda Q^T Q \vec{y}$$

Recordemos que en la matriz ortogonal Q se cumple que $Q^T = Q^{-1}$. Por lo tanto,

$$\vec{x}^T A \vec{x} = \vec{y}^T Q^{-1} Q \Lambda Q^{-1} Q \vec{y} = \vec{y}^T \Lambda \vec{y}$$

Como se puede apreciar, hemos obtenido una **nueva forma cuadrática que no altera a la inicial**, ya que $\vec{x}^T A \vec{x} = \vec{y}^T \Lambda \vec{y}$. Lo que sí cambiamos, fueron los ejes que ahora están dados por \vec{y} en vez de \vec{x} . Este proceso se conoce como el **Teorema del Eje Principal**.

La fortaleza del teorema del eje principal es que obtenemos una forma cuadrática simplificada de la original o, en otras palabras, más sencilla, porque **elimina el término de producto cruzado**. Por ejemplo, digamos que

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Entonces,

$$\vec{y}^T \Lambda \vec{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \lambda_1 y_1^2 + \lambda_2 y_2^2$$

Por otra parte, es fácil conocer al \vec{y} . Recordemos que

$$\vec{x} = Q \vec{y}$$

Multipliquemos a la izquierda por Q^T en ambos lados de la ecuación.

$$Q^T \vec{x} = Q^T Q \vec{y}$$

Como $Q^T = Q^{-1}$, entonces $Q^T Q = I$. En consecuencia

$$\vec{y} = Q^T \vec{x}$$

Esta igualdad nos muestra algo muy relevante con respecto al teorema del eje principal y es que los **nuevos ejes** están **en dirección del los eigenvectores de A** , mientras que sus **longitudes** dependen de los **eigenvalores** de dicha matriz.

Ahora, establezcamos que

$$\vec{y}^T \Lambda \vec{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 = k$$

donde $k = \text{constante}$.

Luego, multipliquemos la ecuación por $1/k$.

$$\frac{\lambda_1}{k} y_1^2 + \frac{\lambda_2}{k} y_2^2 = 1$$

Cuando la matriz A es **positiva definida**, en la ecuación se arriba se cumple que:

$$\frac{1}{\lambda_2} = \frac{\lambda_1}{k}; \quad \frac{1}{\lambda_1} = \frac{\lambda_2}{k}$$

Por lo tanto,

$$\frac{y_1^2}{\lambda_2} + \frac{y_2^2}{\lambda_1} = 1$$

Esta es la **ecuación canónica de la elipse**, pero para los **nuevos ejes dados por \vec{y}** .

Como tenemos la ecuación canónica de la elipse, podemos buscar los valores de a y b para

obtener la altura y ancho de esta figura. En la que obtuvimos, tenemos que:

$$\frac{1}{\lambda_2} = \frac{\lambda_1}{k}; \quad \frac{1}{\lambda_1} = \frac{\lambda_2}{k}$$

En otras palabras, podemos decir que tenemos los recíprocos de los cuadrados de a y b en $1/\lambda_2$ y $1/\lambda_1$. Por lo tanto, las ecuaciones de arriba podemos, en primer lugar, elevarlas por -1 para conocer dichos valores

$$\lambda_2 = \frac{k}{\lambda_1}; \quad \lambda_1 = \frac{k}{\lambda_2}$$

y, luego, tomar la raíz cuadrada en ambas para obtener la mitad de los ejes de la elipse.

$$a = \sqrt{\lambda_2} = \sqrt{\frac{k}{\lambda_1}}; \quad b = \sqrt{\lambda_1} = \sqrt{\frac{k}{\lambda_2}}$$

Por lo tanto, el **ancho** y la **altura** de la elipse canónica en los ejes dados por \vec{y} son:

$$\text{Ancho} = 2a = 2\sqrt{\frac{k}{\lambda_1}}; \quad \text{Altura} = 2b = 2\sqrt{\frac{k}{\lambda_2}}$$

Puesto que los ejes de la elipse de medidas a y b están en dirección de los eigenvectores de la matriz A , podemos buscar los puntos en donde los primeros cortan a esta figura como los componentes de los vectores \vec{a} y \vec{b} que corresponden a los eigenvectores escalados.

$$\vec{a} = c\vec{w}_1 \quad \vec{b} = d\vec{w}_2$$

donde $c, d = \text{constantes}$ y $\vec{w}_1, \vec{w}_2 = \text{eigenvectores de } A$.

Además, debe cumplirse que:

$$||\vec{a}|| = a \quad ||\vec{b}|| = b$$

Como \vec{a} y \vec{b} están en dirección de \vec{w}_1 y \vec{w}_2 , respectivamente, la magnitud de los primeros debe ser igual a la longitud de los segundos aumentados por la misma constante que los escala.

$$||\vec{a}|| = c \cdot ||\vec{w}_1|| \quad ||\vec{b}|| = d \cdot ||\vec{w}_2||$$

Esto nos permite conocer el valor de las constantes c y d .

$$c = \frac{||\vec{a}||}{||\vec{w}_1||} \quad d = \frac{||\vec{b}||}{||\vec{w}_2||}$$

Debido a que $||\vec{a}|| = a$ y $||\vec{b}|| = b$, entonces:

$$c = \frac{a}{||\vec{w}_1||} \quad d = \frac{b}{||\vec{w}_2||}$$

También, sabemos que $a = \sqrt{k/\lambda_1}$ y $b = \sqrt{k/\lambda_2}$.

$$c = \frac{\sqrt{k/\lambda_1}}{||\vec{w}_1||} \quad d = \frac{\sqrt{k/\lambda_2}}{||\vec{w}_2||}$$

En consecuencia,

$$\vec{a} = \frac{\sqrt{k/\lambda_1}}{||\vec{w}_1||} \cdot \vec{w}_1 \quad \vec{b} = \frac{\sqrt{k/\lambda_2}}{||\vec{w}_2||} \cdot \vec{w}_2$$

donde los componentes de \vec{a} y \vec{b} son los puntos donde la mitad de los ejes principales cortan a la elipse.

Además, se cumple que la **magnitud del eje más largo** de la elipse está determinado por el **eigenvalor de A más chico**, mientras que la **longitud del eje más corto** depende del **eigenvalor más grande** de la misma matriz.

Ejemplo 1. Identifique los ejes principales de la elipse dada por la ecuación $5x^2 - 4xy + 8y^2 = 16$.

Solución. Comencemos graficando la forma cuadrática para tener una idea visual del nivel de inclinación de la elipse.

Es más rápido graficar una forma cuadrática por medio de un software, pero si en éste no podemos trabajar de forma directa con la ecuación original, una alternativa es pasarla a coordenadas polares y, luego, parametrizar sus variables iniciales en términos de un ángulo θ .

En general, si una forma cuadrática es de la forma

$$ax^2 + 2bxy + cy^2 = k,$$

al establecer que $x(\theta) = r(\theta) \cos(\theta)$ e $y(\theta) = r(\theta) \sin(\theta)$, donde θ es un ángulo, podemos

reescribirla como:

$$a(r(\theta) \cos(\theta))^2 + 2b(r(\theta) \cos(\theta))(r(\theta) \sin(\theta)) + c(r(\theta) \sin(\theta))^2 = k$$

Es posible simplificar un poco la expresión de la izquierda de la ecuación por el hecho de que $\sin(2\theta) = 2 \cos(\theta) \sin(\theta)$.

$$ar^2(\theta) \cos^2(\theta) + 2br^2(\theta) \sin(2\theta) + r^2(\theta) \sin^2(\theta) = k$$

Al factorizar y al despejar a $r(\theta)$, obtenemos la forma cuadrática como ecuación polar.

$$r(\theta) = \sqrt{\frac{k}{a \cos^2(\theta) + b \sin(2\theta) + c \sin^2(\theta)}}$$

Por lo tanto, podemos definir que:

$$x(\theta) = \left(\sqrt{\frac{k}{a \cos^2(\theta) + b \sin(2\theta) + c \sin^2(\theta)}} \right) \cdot \cos(\theta)$$

$$y(\theta) = \left(\sqrt{\frac{k}{a \cos^2(\theta) + b \sin(2\theta) + c \sin^2(\theta)}} \right) \cdot \sin(\theta)$$

Al reemplazar en $x(\theta)$ e $y(\theta)$ usando los valores de la ecuación $5x^2 - 4xy + 8y^2 = 16$, obtenemos la gráfica de dicha forma cuadrática.



Ahora calculemos los eigenvalores y eigenvectores de la matriz A , que corresponde a:

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}$$

Como vemos, A es simétrica y sus subdeterminantes $\det(5)$, $\det(A) > 0$, por tanto podemos dar por sentado de que es positiva definida⁶ y, en consecuencia, que sus eigenvalores son distintos, reales y positivos, mientras que sus eigenvectores son ortogonales.

Obtengamos los eigenvalores de A usando la ecuación característica.

$$\begin{aligned} \det(A - \lambda I) &= 0 \\ \begin{vmatrix} 5 - \lambda & -2 \\ -2 & 8 - \lambda \end{vmatrix} &= 0 \\ \lambda^2 - 13\lambda + 36 &= 0 \\ (\lambda_1 - 4)(\lambda_2 - 9) &= 0 \end{aligned}$$

Por lo tanto, los eigenvalores de A son:

$$\lambda_1 = 4 \quad \lambda_2 = 9$$

Luego, calculemos los eigenvectores \vec{w}_i para cada λ_i . Para el caso de $\lambda_1 = 4$ tenemos que

$$\begin{aligned} (A - \lambda_1 I)\vec{w}_1 &= \vec{0} \\ \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ w_1 - 2 &= 0 \\ w_1 &= 2 \end{aligned}$$

Por lo tanto, el eigenvector asociado a $\lambda_1 = 4$ es:

$$\vec{w}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

⁶Aunque esto lo deducimos antes al saber que la gráfica de la forma cuadrática es una elipse.

En cuanto al eigenvalor $\lambda_2 = 9$ de la matriz A , su eigenvector \vec{w}_2 podemos obtenerlo como:

$$\begin{aligned}(A - \lambda_2 I)\vec{w}_2 &= \vec{0} \\ \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} -4 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ -4w_1 - 2 &= 0 \\ w_1 &= -\frac{1}{2}\end{aligned}$$

Por lo tanto,

$$\vec{w}_2 = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}$$

Luego, normalicemos los eigenvectores \vec{w}_1 y \vec{w}_2 , los cuales denotaremos como \vec{q}_1 y \vec{q}_2 , respectivamente.

$$\begin{aligned}\vec{q}_1 &= \frac{1}{\|\vec{w}_1\|} \cdot \vec{w}_1 = \frac{1}{\sqrt{5}} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \\ \vec{q}_2 &= \frac{1}{\|\vec{w}_2\|} \cdot \vec{w}_2 = \frac{1}{\sqrt{5}/2} \cdot \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}\end{aligned}$$

A partir de esta información, calculemos las longitudes de los ejes principales de la elipse.

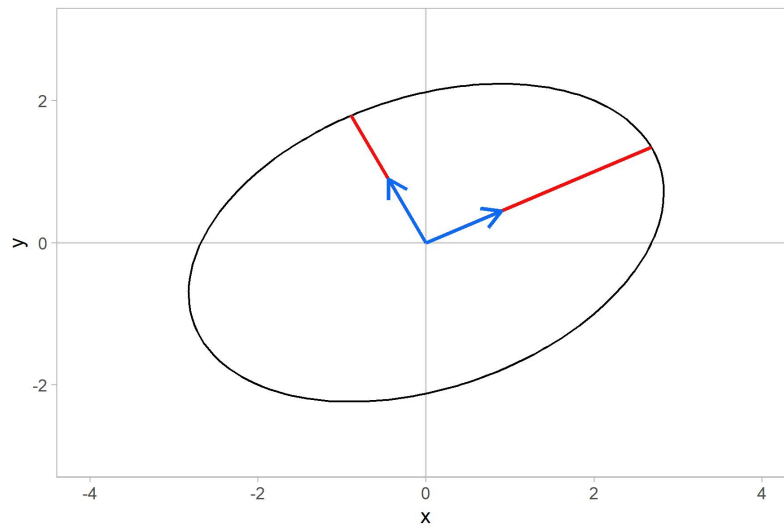
$$a = \sqrt{\frac{36}{4}} = 3 \qquad b = \sqrt{\frac{36}{9}} = 2$$

Por lo tanto, el ancho de la elipse es $2a = 6$ y su altura $2b = 4$.

Finalmente, veamos en qué puntos los ejes principales cortan a la elipse. Para ello, usemos los eigenvectores de A normalizados, ya que sabemos que sus magnitudes son iguales a 1.

$$\begin{aligned}\vec{a} &= \frac{3}{1} \cdot \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ 3/\sqrt{5} \end{bmatrix} \\ \vec{b} &= \frac{2}{1} \cdot \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -2/\sqrt{5} \\ 4/\sqrt{5} \end{bmatrix}\end{aligned}$$

A continuación tenemos la gráfica de la elipse con los eigenvectores de la matriz A normalizados (en azul) y sus ejes principales (en rojo).



Por otra parte, si establecemos que $\vec{y} = [y_1 \ y_2]^T$ corresponde a los nuevos ejes de la elipse, podemos reescribir la forma cuadrática como una ecuación canónica de esta figura.

$$\frac{y_1^2}{9} + \frac{y_2^2}{4} = 1$$

La elipse canónica de esta ecuación podemos verla en la siguiente gráfica, la cual es la misma a la original, pero en este nuevo sistema sus ejes cambiaron tal que fuesen ortogonales.

