

Formas cuadráticas

Fernando Chamizo

23 de abril de 2019

1. Definición

Las formas bilineales en física aparecen a menudo evaluadas con los dos argumentos iguales. Por ejemplo, el tensor de inercia es una forma bilineal $I = I(\vec{x}, \vec{y})$ pero gran parte de su interés radica en que $\frac{1}{2}I(\vec{\omega}, \vec{\omega})$ da la energía de rotación cuando la velocidad angular es $\vec{\omega}$.

Con esta motivación se define una *forma cuadrática* como una aplicación $Q : V \rightarrow \mathbb{K}$ del tipo $Q(\vec{x}) = \varphi(\vec{x}, \vec{x})$ con $\varphi : V \times V \rightarrow \mathbb{K}$ una forma bilineal. Aquí V es un espacio vectorial de dimensión finita sobre $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$ aunque este último caso es de interés menor y en casi todo el tema nos restringiremos a $\mathbb{K} = \mathbb{R}$.

Sin avisarlo cada vez identificaremos vectores y sus coordenadas con el abuso de notación obvio. Es decir, supondremos implícitamente que hay una base fijada.

Ejemplo. El producto escalar usual sobre \mathbb{R}^3 define la forma cuadrática $Q(\vec{x}) = x_1^2 + x_2^2 + x_3^2$ (usando coordenadas en la base canónica). Esta misma fórmula para Q también define una forma cuadrática sobre \mathbb{C} pero no proviene del producto escalar usual de \mathbb{C}^3 porque este da una forma sesquilineal, no bilineal.

En principio las formas cuadráticas escapan del negociado de este curso donde todos los objetos han sido lineales en una o dos variables. El motivo por el que se incorporan a los temarios de álgebra lineal es que se estudian con herramientas de esta área. La razón última es que hay una biyección que preserva la estructura entre formas cuadráticas y ciertas formas bilineales.

Ejemplo. Las siguientes formas bilineales $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$

$$(1) \quad \varphi_1(\vec{x}, \vec{y}) = \vec{x}^t \begin{pmatrix} 1 & 5 \\ 3 & 7 \end{pmatrix} \vec{y} \quad y \quad \varphi_2(\vec{x}, \vec{y}) = \vec{x}^t \begin{pmatrix} 1 & 1 \\ 7 & 7 \end{pmatrix} \vec{y}$$

cumplen $\varphi_1(\vec{x}, \vec{x}) = \varphi_2(\vec{x}, \vec{x}) = x_1^2 + 8x_1x_2 + 7x_2^2$ y por tanto dan lugar a la misma forma cuadrática.

Para eliminar la redundancia mostrada en el ejemplo, la clave es observar que toda matriz se escribe como suma de una simétrica y otra antisimétrica

$$(2) \quad A = M_s + M_a \quad \text{con} \quad M_s = \frac{A + A^t}{2} \quad y \quad M_a = \frac{A - A^t}{2}.$$

Esto implica que cualquier forma bilineal $\varphi(\vec{x}, \vec{x}) = \vec{x}^t A \vec{x}$ se puede escribir como $\varphi = \varphi_s + \varphi_a$ con φ_s *simétrica*, lo que significa $\varphi_s(\vec{x}, \vec{y}) = \varphi_s(\vec{y}, \vec{x})$, y con φ_a *antisimétrica*, que análogamente significa $\varphi_a(\vec{x}, \vec{y}) = -\varphi_a(\vec{y}, \vec{x})$. Claramente las formas bilineales antisimétricas dan lugar a la forma cuadrática nula y así formas bilineales como las del ejemplo anterior que difieran en una forma antisimétrica dan lugar a la misma forma cuadrática. Es decir, que tenemos una aplicación sobreyectiva de las formas simétricas en las formas cuadráticas. Para ver que es biyectiva hay que comprobar que la forma cuadrática permite recuperar la forma bilineal simétrica de la que proviene. Esto suena un poco extraño porque a fin de cuentas la forma bilineal tiene dos variables y parece contener más información que la forma cuadrática que tiene solo una pero la simetría de algún modo reduce la información a la mitad. Si $Q(\vec{x}) = \varphi(\vec{x}, \vec{x})$, la fórmula para recuperar φ a partir de Q es

$$(3) \quad \varphi(\vec{x}, \vec{y}) = \frac{1}{2}(Q(\vec{x} + \vec{y}) - Q(\vec{x}) - Q(\vec{y})).$$

La prueba se reduce a usar la igualdad $\varphi(\vec{x} + \vec{y}, \vec{x} + \vec{y}) = \varphi(\vec{x}, \vec{x}) + \varphi(\vec{y}, \vec{y}) + \varphi(\vec{x}, \vec{y}) + \varphi(\vec{y}, \vec{x})$. La clave es que la simetría implica que los dos últimos términos son iguales y así se puede despejar $\varphi(\vec{x}, \vec{y})$.

Una vez vista la equivalencia entre formas cuadráticas y formas bilineales simétricas, se definen la *matriz* de una forma cuadrática y el *cambio de base* de una forma cuadrática de manera similar a la conocida para las formas bilineales. Así pues la forma cuadrática de matriz A , necesariamente simétrica, es

$$(4) \quad Q(\vec{x}) = \vec{x}^t A \vec{x} = \sum_i \sum_j a_{ij} x_i x_j.$$

Ejemplo. La matriz de la forma $3x^2 + 5xy + 7y^2$ (de \mathbb{R}^2 en \mathbb{R}) es

$$(5) \quad A = \begin{pmatrix} 3 & 5/2 \\ 5/2 & 7 \end{pmatrix} \quad \text{y en general} \quad \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \quad \text{es la de} \quad ax^2 + bxy + cy^2.$$

La reducción a la mitad de los elementos fuera de la diagonal se debe a que a_{ij} y a_{ji} afectan al mismo producto $x_i x_j$.

Los cambios de base son cambios de variable y al efectuar en $\vec{x}^t A \vec{x}$ el cambio $\vec{x} \mapsto C \vec{x}'$ la matriz se modifica por medio de $A \mapsto C^t A C$, como en el caso bilineal.

Ejemplo. Hallemos la matriz de $Q(\vec{x}) = x_1^2 - x_2^2$, definida en \mathbb{R}^2 con la base canónica, en las bases $B_1 = \{(1, 0)^t, (2, 1)^t\}$ y $B_2 = \{(5/4, 3/4)^t, (3/4, 5/4)^t\}$. Las matrices buscadas son respectivamente

$$(6) \quad \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \quad \text{y} \quad \frac{1}{16} \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

El hecho de que la matriz de Q haya permanecido invariante en la segunda base tiene su interpretación física en que el cambio de base está asociado a una transformación de Lorentz.

2. Diagonalización de formas cuadráticas

Se dice que una forma cuadrática diagonaliza en una base si en dicha base su matriz es diagonal. Consecuentemente es del tipo $\sum_i \lambda_i x_i^2$.

Supongamos que V es un espacio euclídeo, en particular sobre $\mathbb{K} = \mathbb{R}$. Utilizando una base ortonormal siempre podemos suponer que estamos en \mathbb{R}^n con el producto escalar usual. En esta situación, una matriz simétrica A corresponde en la base canónica a una aplicación autoadjunta y el teorema espectral asegura que se diagonaliza en una base ortonormal. Los cambios entre base ortonormales vienen dados por matrices ortogonales y se tiene

$$(7) \quad C^{-1}AC = D \quad \text{con } D \text{ diagonal y } C^tC = I.$$

Esto implica $C^{-1}AC = C^tAC$ y por tanto una forma cuadrática se diagonaliza simplemente pasando a una base ortonormal de autovectores. Recordemos que autovalores distintos de una aplicación autoadjunta corresponden a autovectores ortogonales, así que basta con normalizar los autovectores excepto en el caso de autovalores múltiples en que tendremos que forzar a que nuestra elección sea ortogonal antes de normalizar.

Ejemplo. Para diagonalizar la forma cuadrática $x^2 + 2xy + 3y^2$ en una base ortonormal, escribimos su matriz y calculamos los autovalores,

$$(8) \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \quad |A - \lambda I| = 0 \quad \Rightarrow \quad \lambda_{\pm} = 2 \pm \sqrt{2}.$$

Los autovectores respectivos son $\vec{w}_{\pm} = (1, 1 \pm \sqrt{2})^t$. Necesariamente son ortogonales y al normalizarlos se obtiene $\vec{v}_1 = \vec{w}_+ / \|\vec{w}_+\|$, $\vec{v}_2 = \vec{w}_- / \|\vec{w}_-\|$. Se sigue que $\{\vec{v}_1, \vec{v}_2\}$ es una base ortonormal en la que la matriz de la forma cuadrática es la matriz diagonal con elementos λ_+ y λ_- . Si no normalizásemos, la matriz seguiría siendo diagonal pero ya no se obtendrían los autovalores como elementos.

Ejemplo. Consideremos en \mathbb{R}^3 la forma $8x^2 + 5y^2 + 5z^2 - 4xy + 4xz + 8yz$. Su matriz y polinomio característico son

$$(9) \quad A = \begin{pmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{pmatrix}, \quad |A - \lambda I| = -\lambda(\lambda - 9)^2.$$

Buscamos una base ortonormal de autovectores $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. El autovalor $\lambda = 0$ conduce al autovector $(1, 2, -2)^t$ que normalizado da $\vec{v}_1 = \frac{1}{3}(1, 2, -2)^t$. Seguro que los autovectores correspondientes a $\lambda = 9$ forman un subespacio de dimensión dos ortogonal a \vec{v}_1 pero una elección al azar de una base suya no nos dará en general vectores ortogonales entre sí. Digamos que tomamos $\vec{w}_1 = (2, 1, 2)^t$ y $\vec{w}_2 = (2, 0, 1)^t$. Podemos forzar a que \vec{w}_2 sea ortogonal con el proceso de Gram-Schmidt. Es decir, sustituyéndolo por $\vec{w}_2 - (\vec{w}_1 \cdot \vec{w}_2)\vec{w}_1 / \|\vec{w}_1\|^2$ que es $\frac{1}{3}(2, -2, -1)^t$. Casualmente está ya normalizado, mientras que es necesario normalizar \vec{w}_1 . En definitiva, $\vec{v}_2 = \frac{1}{3}(2, 1, 2)^t$ y $\vec{v}_3 = \frac{1}{3}(2, -2, -1)^t$.

Para que sea posible calcular autovalores y autovectores explícitamente, el problema tiene que estar “preparado”, ya que es muy raro que tengamos fórmulas explícitas para las soluciones de una ecuación algebraica elegida al azar. Si sacrificamos la condición de que la diagonalización se produzca en una base ortonormal, hay un algoritmo que permite diagonalizar de una manera fácil de implementar en un ordenador y que da resultados totalmente explícitos. Este algoritmo se llama *método de Gauss*, igual que el de resolución de ecuaciones lineales, y a pesar de tener un nombre propio es una gran obviedad que esencialmente se reduce a completar cuadrados.

Ejemplo. Para diagonalizar la forma cuadrática $x^2 + 2xy + 3y^2$ necesitamos eliminar el término $2xy$, de esta forma agrupamos $(x + y)^2 + 2y^2$. ahora definiendo unas nuevas variables $x' = x + y$, $y' = y$ tendremos la forma diagonalizada $(x')^2 + 2(y')^2$. Este cambio de variable es un cambio de base

$$(10) \quad \vec{x}' = C\vec{x} \quad \text{con} \quad C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{y se tiene} \quad C^t \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} C = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}.$$

El cambio C va en el sentido contrario al que se obtendría calculando autovectores y para hallar las coordenadas de los vectores de la base $\{\vec{v}_1, \vec{v}_2\}$ en la que diagonaliza, habría que hallar las columnas de C^{-1} . Concretamente, $\vec{v}_1 = (1, 0)^t$ y $\vec{v}_2 = (-1, 1)^t$. Esta base no es ortogonal.

Como receta general podemos usar

$$(11) \quad a_{kk}x_k^2 + \sum_{j=k}^n a_{kj}x_kx_j = a_{kk} \left(x_k + \sum_{j=k+1}^n \frac{a_{kj}}{2a_{kk}}x_j \right)^2 - a_{kk} \left(\sum_{j=k+1}^n \frac{a_{kj}}{2a_{kk}}x_j \right)^2$$

para eliminar primero los términos cruzados x_1x_j , $j > 1$; después los términos cruzados x_2x_j , $j > 2$, etc. Cada término entre paréntesis dará lugar a una nueva variable y a una fila en la matriz C . Lo único que puede impedir que continuemos este proceso es que $a_{kk} = 0$. En esa situación reordenar las variables salva este obstáculo siempre que haya algún a_{kk} restante no nulo.

Ejemplo. Consideremos la forma cuadrática $Q = 2x^2 + 12xy + 18y^2 - 4xz - 6yz + 3z^2$ definida en \mathbb{R}^3 o en \mathbb{C}^3 . Al aplicar (11) en x , la primera variable,

$$(12) \quad Q = (2(x')^2 - 2(3y - z)^2) + 18y^2 - 6yz + 3z^2 \quad \text{con} \quad x' = x + 3y - z.$$

Al operar $Q = 2(x')^2 + 6yz + z^2$ y como el término y^2 ha desaparecido no es posible completar cuadrados con (11) en y . Sin embargo, si decidimos que z es la segunda variable no hay problema en continuar para obtener

$$(13) \quad Q = 2(x')^2 + ((z')^2 - 9y^2) \quad \text{con} \quad z' = z + 3y.$$

De esta forma hemos llegado a la diagonalización

$$(14) \quad Q = 2(x')^2 - 9(y')^2 + (z')^2 \quad \text{con} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

El único caso extremo que se escapa se produce cuando entre las variables restantes no aparece ninguna elevada al cuadrado. En ese caso hay un truco muy simple que consiste en que $x_j x_k$ se convierte en $\tilde{x}_j^2 - \tilde{x}_k^2$ bajo el cambio artificial $x_j = \tilde{x}_j + \tilde{x}_k$, $x_k = \tilde{x}_j - \tilde{x}_k$.

En definitiva, el *método de Gauss* consiste en aplicar (11) siempre que sea posible, intercalando este cambio artificial si no lo fuera.

Una observación final es que el método de Gauss, a diferencia del basado en autovalores y autovectores, no se limita a $\mathbb{K} = \mathbb{R}$, sirve igualmente para $\mathbb{K} = \mathbb{C}$ porque (11) es una identidad válida en cualquier cuerpo (en rigor en el que $1 + 1 \neq 0$). Por ejemplo, si todos los coeficientes son racionales todos los cálculos serán con números racionales.

Ejemplo. Consideremos la forma cuadrática en \mathbb{R}^3

$$(15) \quad Q(\vec{x}) = \vec{x}^t A \vec{x} \quad \text{donde} \quad A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Escribiendo $\vec{x} = (x, y, z)$ se tiene $Q = 2xy - 2xz$ con lo que no hay términos cuadráticos en ninguna de las variables. Cambiamos entonces $x = \tilde{x} + \tilde{y}$, $y = \tilde{x} - \tilde{y}$, renombrando $z = \tilde{z}$ para mayor simetría notacional. Con estas coordenadas $Q = 2\tilde{x}^2 - 2\tilde{y}^2 - 2\tilde{x}\tilde{z} - 2\tilde{y}\tilde{z}$. Ahora ya estamos en condiciones de completar cuadrados con (11) en \tilde{x} y se sigue

$$(16) \quad Q = \left(2(x')^2 - 2\left(\frac{1}{2}\tilde{z}\right)^2\right) - 2\tilde{y}^2 - 2\tilde{y}\tilde{z} \quad \text{con} \quad x' = \tilde{x} - \frac{1}{2}\tilde{z}.$$

Ahora continuamos con $Q = 2(x')^2 - 2\tilde{y}^2 - 2\tilde{y}\tilde{z} - \tilde{z}^2/2$ completando cuadrados en \tilde{y} (también se podría en \tilde{z}) para obtener

$$(17) \quad Q = \left(2(x')^2 + (-2(y')^2 + 2\left(\frac{1}{2}\tilde{z}\right)^2)\right) - \frac{1}{2}\tilde{z}^2 \quad \text{con} \quad y' = \tilde{y} + \frac{1}{2}\tilde{z}.$$

En definitiva, hemos diagonalizado a $Q = 2(x')^2 - 2(y')^2$ con $x' = \tilde{x} - \tilde{z}/2$, $y' = \tilde{x} + \tilde{z}/2$, $z' = \tilde{z}$. En términos de las variables originales, como $\tilde{x} = \frac{1}{2}(x + y)$, $\tilde{y} = \frac{1}{2}(x - y)$, se tiene $x' = \frac{1}{2}(x + y - z)$, $y' = \frac{1}{2}(x - y + z)$, $z' = z$. Matricialmente,

$$(18) \quad A = C^t \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} C \quad \text{con} \quad C = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

3. Ley de inercia y clasificación

Ciertamente no hay una manera única de diagonalizar una forma cuadrática. Esto contrasta con lo que ocurría con aplicaciones lineales. Para estas uno está obligado a escoger una base

de autovectores y cuando los autovalores son distintos la única libertad es encoger o alargar los vectores. Sin embargo diferentes métodos al diagonalizar una forma cuadrática llevan habitualmente a bases que no guardan relaciones sencillas. Una pregunta natural es si hay algo que deba respetarse, permanecer invariante, con toda esta libertad.

Digamos que en dimensión n tenemos $C^t AC = D$ con D diagonal. Si C proviene de un cambio de base, C^t y C son no singulares. Por tanto $C^t A$ tiene el mismo rango que A y lo mismo ocurre con $C^t AC$ (aquí se está usando que el rango por filas y columnas coincide). La conclusión es que el número de ceros en la diagonal de D es $n - \text{rg}(A)$ y por tanto no depende del método utilizado para diagonalizar.

Para $\mathbb{K} = \mathbb{C}$ de aquí se deduce una clasificación sencilla de las formas cuadráticas. Si la matriz de una forma cuadrática Q tiene rango m entonces diagonalizará como $\alpha_1 x_1^2 + \dots + \alpha_m x_m^2$ con $\alpha_j \neq 0$. Ahora bien, sobre los complejos es lícito el cambio $x_j \mapsto x_j / \sqrt{\alpha_j}$ y así toda forma de rango m con un cambio de base se reduce a $x_1^2 + \dots + x_m^2$.

La situación para $\mathbb{K} = \mathbb{R}$ no es tan obvia. Con cambios $x_j \mapsto x_j / \sqrt{|\alpha_j|}$ los coeficientes se pueden reducir a ± 1 pero en principio no está claro si $x^2 - y^2 - z^2$ podría transformarse en $x^2 + y^2 - z^2$. Lo que sí está claro es que $x^2 + y^2 + z^2$ no puede transformarse en $-x^2 - y^2 - z^2$ porque la primera toma valores positivos excepto para $\vec{0}$ mientras que la segunda los toma negativos.

Dada una forma cuadrática Q definida sobre un espacio vectorial de dimensión n sobre $\mathbb{K} = \mathbb{R}$, que siempre podemos suponer \mathbb{R}^n , si se diagonaliza adquiriendo una matriz D , se define su *signatura* como la terna (n_0, n_+, n_-) donde n_0 , n_+ y n_- son el número de elementos nulos, positivos y negativos en la diagonal de D . Evidentemente $n = n_0 + n_+ + n_-$.

La *ley de inercia de formas cuadráticas* afirma que la signatura está bien definida, solo depende de la forma cuadrática y no del método utilizado para diagonalizar. Así pues $x^2 - y^2 - z^2$ que tiene signatura $(0, 1, 2)$ como forma cuadrática $\mathbb{R}^3 \rightarrow \mathbb{R}$, no puede transformarse en $x^2 + y^2 - z^2$ que tiene signatura $(0, 2, 1)$.

A veces cuando el rango es máximo se suprime el primer número en la signatura. Esto es, se abrevia $(0, n_+, n_-)$ como (n_+, n_-) . En física la signatura $(1, 3)$ desempeña un papel destacado. En matemáticas hay poco acuerdo en la definición de signatura y algunos autores ordenan sus elementos de otra forma e incluso la reducen a solo un número.

La prueba de la ley de inercia de formas cuadráticas es solo una pequeña variante del argumento sobre la positividad de $x^2 + y^2 + z^2$. Digamos que una forma cuadrática de rango m tuviera dos signaturas, dependiendo del método usado al diagonalizar. Necesariamente serían de la forma $(n - m, k, m - k)$ y $(n - m, k', m - k')$ porque $n_0 = n - \text{rg}(A)$ y $n = n_0 + n_+ + n_-$. Así se tendría

$$(19) \quad Q = \alpha_1 x_1^2 + \dots + \alpha_k x_k^2 - \alpha_{k+1} x_{k+1}^2 - \dots - \alpha_m x_m^2 \quad \text{en una base } B = \{\vec{v}_1, \dots, \vec{v}_n\}$$

con $\alpha_j > 0$ y también, para $\beta_j > 0$,

$$(20) \quad Q = \beta_1 x_1^2 + \dots + \beta_{k'} x_{k'}^2 - \beta_{k'+1} x_{k'+1}^2 - \dots - \beta_m x_m^2 \quad \text{en una base } B' = \{\vec{v}'_1, \dots, \vec{v}'_n\}.$$

Intercambiando los nombres de B y B' si es necesario, siempre se puede suponer $k > k'$ si las signaturas no son iguales. Si V es el subespacio generado por $\{\vec{v}_1, \dots, \vec{v}_k\}$ entonces $Q(\vec{x}) > 0$ para cualquier $\vec{x} \in V - \{\vec{0}\}$ y si W es el subespacio generado por $\{\vec{v}'_{k'+1}, \dots, \vec{v}'_n\}$ entonces $Q(\vec{x}) \leq 0$ para cualquier $\vec{x} \in W$. Por otro lado $\dim V + \dim W = k + n - k' > n$ y esto implica $V \cap W \neq \{\vec{0}\}$ por la fórmula de Grassman. Esto lleva a una contradicción porque para $\vec{x} \in V \cap W - \{\vec{0}\}$, el valor de $Q(\vec{x})$ debe ser simultáneamente positivo y no negativo.

La conclusión de este análisis es que para $\mathbb{K} = \mathbb{R}$ una forma cuadrática se puede transformar en otra por un cambio de base si y solo si tienen la misma signatura.

Las diferentes posibilidades para la signatura llevan a la siguiente clasificación donde k, l y m indican enteros positivos:

$$(21) \quad \left\{ \begin{array}{ll} (0, l, 0) & \rightarrow \textit{definida positiva} : \quad Q(\vec{x}) > 0 \text{ excepto } Q(\vec{0}) = 0. \\ (0, 0, m) & \rightarrow \textit{definida negativa} : \quad Q(\vec{x}) < 0 \text{ excepto } Q(\vec{0}) = 0. \\ (k, l, 0) & \rightarrow \textit{semidefinida positiva} : \quad Q(\vec{x}) \geq 0 \text{ con } Q(\vec{x}) = 0 \text{ para algún } \vec{x} \neq \vec{0}. \\ (k, 0, m) & \rightarrow \textit{semidefinida negativa} : \quad Q(\vec{x}) \leq 0 \text{ con } Q(\vec{x}) = 0 \text{ para algún } \vec{x} \neq \vec{0}. \\ \left. \begin{array}{l} (0, l, m) \\ (k, l, m) \end{array} \right\} & \rightarrow \textit{indefinida} : \quad \text{Existen } \vec{x} \text{ y } \vec{x}' \text{ con } Q(\vec{x}) > 0 \text{ y } Q(\vec{x}') < 0. \end{array} \right.$$

En general a las formas cuadráticas que tienen $n_0 > 0$ se les llama *degeneradas*, lo que equivale a que tengan una matriz singular. Así se podrían distinguir los dos tipos de formas indefinidas diciendo que las primeras son indefinidas no degeneradas y las segundas indefinidas degeneradas.

Recordemos que para formas no degeneradas el criterio de Sylvester permite decidir el tipo de forma cuadrática: Es definida positiva si y solo si los *menores angulares* de la matriz de Q son todos positivos, es definida negativa si y solo si se cumple la condición para el negativo de la matriz y es indefinida en el resto de los casos. En cuanto a la eficiencia computacional este criterio es inabordable en cuanto la dimensión es un poco grande y el método de Gauss se vuelve mucho más adecuado además de incorporar también los casos semidefinidos.

4. Diagonalización simultánea de dos formas cuadráticas

La rigidez que impone la diagonalización de aplicaciones lineales hace casi imposible que dos de ellas se diagonalicen en la misma bases porque deberían coincidir sus autovectores. De hecho en el caso autoadjunto (que asegura la existencia de diagonalización) eso es equivalente a la conmutación lo cual tiene implicaciones interesantes en física.

Las formas cuadráticas permiten mucha más libertad y cabe preguntarse si es posible diagonalizar simultáneamente dos de ellas en un espacio vectorial real. La respuesta no es siempre afirmativa pero sí lo es en el caso real cuando una de las formas es definida positiva.

Es muy natural preguntarse por qué estamos interesados en este problema tan peculiar. La motivación es principalmente física. Veámosla a continuación sin mucho detalle.

La energía cinética de un sistema mecánico es una forma cuadrática definida positiva en las velocidades de las partículas, digamos $Q_k(\vec{v})$. Aproximando el potencial cerca de un punto crítico \vec{x}_0 por medio de Taylor de orden 2 se tiene que la energía potencial se aproxima por una constante más una forma cuadrática $Q_p(\vec{x} - \vec{x}_0)$. Con un cambio de los orígenes de coordenadas y energías (lo que no afecta a la mecánica) se puede suponer $\vec{x}_0 = \vec{0}$ y que la energía potencial es una forma cuadrática $Q_p(\vec{x})$ en las posiciones. Una diagonalización simultánea de Q_k y Q_p implica que en una base adecuada la energía es de la forma $\sum_{i=1}^n (\mu_i v_i^2 + \lambda_i x_i^2)$. Esto es proporcional a la energía que correspondería a n partículas desacopladas de masas μ_i sometidas a fuerzas unidimensionales $-\lambda_i x_i$ porque $\mu_i \frac{d^2 x_i}{dt^2} = -\lambda_i x_i$ implica $\frac{d}{dt}(\mu_i v_i^2 + \lambda_i x_i^2) = 0$ con $v_i = \frac{dx_i}{dt}$. Si los λ_i son positivos entonces son n osciladores armónicos unidimensionales desacoplados $x_i = a_i \cos(\omega_i t) + b_i \sin(\omega_i t)$ con frecuencias angulares $\omega_i = \sqrt{\lambda_i / \mu_i}$. Al deshacer el cambio que diagonaliza simultáneamente deducimos que cerca de un punto crítico del potencial en el que $\lambda_i > 0$ (lo que implica que alcanza un mínimo local) el sistema muestra en primera aproximación unas oscilaciones características dadas por

$$(22) \quad \vec{x} = \sum_{i=1}^n (a_i \cos(\omega_i t) + b_i \sin(\omega_i t)) \vec{u}_i$$

para ciertos vectores constantes \vec{u}_i que se dice que dan los *modos normales de vibración* mientras que a las ω_i se les llama *frecuencias naturales*. Si algún λ_i fuera negativo, entonces habría un comportamiento inestable exponencial.

Para hacer la teoría matemática, digamos que las formas cuadráticas son $Q_1(\vec{x}) = \vec{x}^t A \vec{x}$ y $Q_2(\vec{x}) = \vec{x}^t B \vec{x}$ con Q_2 definida positiva. Veamos primero que verdaderamente existe una diagonalización simultánea. Por la clasificación de las formas cuadráticas, debe existir una matriz C_2 tal que $C_2^t B C_2 = I$. Como $C_2^t A C_2$ es simétrica, diagonaliza en base ortonormal y existe C_1 matriz ortogonal tal que $C_1^t C_2^t A C_2 C_1$ es diagonal. Es decir, que Q_1 diagonaliza con el cambio de base $\vec{x} \mapsto C_2 C_1 \vec{x}$ pero este cambio también diagonaliza Q_2 , de hecho hace que su matriz sea la identidad porque $C_1^t C_2^t B C_2 C_1 = C_1^t I C_1 = I$, ya que C_1 es ortogonal.

La demostración anterior es constructiva pero laboriosa en la práctica. Hay un atajo para calcular $C = C_2 C_1$ sin necesidad de hallar C_1 y C_2 por separado. Para ello notemos como C cumple $C^t A C = D$ (diagonal) y $C^t B C = I$ entonces $|D - \lambda I| = 0$ tiene las mismas soluciones que $|A - \lambda B| = 0$ y así resolviendo esta variante de la ecuación característica se obtienen los elementos $\lambda_1, \dots, \lambda_n$ de la diagonal de D (relacionados con el cuadrado de las frecuencias). Por otro lado, los vectores columna \vec{v}_j de C cumplen

$$(23) \quad C^t A \vec{v}_j = C^t A C \vec{e}_j = \lambda_j \vec{e}_j \quad \text{y} \quad C^t B \vec{v}_j = C^t B C \vec{e}_j = I \vec{e}_j = \vec{e}_j.$$

Entonces son soluciones de $C^t(A - \lambda_j B)\vec{x} = \vec{0}$. Además $C^t BC = I$ impone que estas columnas son ortonormales con respecto al producto escalar $\langle \vec{x}, \vec{y} \rangle = \vec{x}^t B \vec{y}$. Uniendo todas estas piezas, se tiene el siguiente algoritmo:

1. Resolver la ecuación $|A - \lambda B| = 0$. Las soluciones $\lambda_1, \dots, \lambda_n$ son los elementos de la diagonal de D (repetidos con su multiplicidad).
2. Hallar una base $\{\vec{x}_1, \dots, \vec{x}_n\}$ formada por vectores de los núcleos de $A - \lambda_j B$.
3. Ortonormalizar \vec{x}_j con el producto escalar $\langle \vec{x}, \vec{y} \rangle = \vec{x}^t B \vec{y}$. Si la multiplicidad es uno esto se reduce a dividir entre $(\vec{x}_j^t B \vec{x}_j)^{1/2}$.

Los vectores obtenidos son las columnas del cambio de base C buscado. Nótese el paralelismo con el algoritmo para diagonalizar en base ortonormal $\vec{x}^t A \vec{x}$. Se vuelven idénticos si $B = I$.

Ejemplo. Diagonalicemos simultáneamente las formas cuadráticas en \mathbb{R}^2 dadas por $Q_1 = -4xy$ y $Q_2 = x^2 - 2xy + 4y^2$ que tienen por matrices

$$(24) \quad A = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}.$$

La segunda forma cuadrática es definida positiva por el criterio de Sylvester o simplemente escribiéndola como $(x-y)^2 + 3y^2$. Al resolver $|A - \lambda B| = 0$ se obtienen $\lambda_1 = -2$ y $\lambda_2 = 2/3$. Ahora escogemos elementos en los núcleos de $A - \lambda_j B$:

$$(25) \quad (A - \lambda_1 B)\vec{x}_1 = \vec{0} \quad \text{con} \quad \vec{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{y} \quad (A - \lambda_2 B)\vec{x}_2 = \vec{0} \quad \text{con} \quad \vec{x}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Como los núcleos son de dimensión 1, la única libertad es escoger múltiplos de estos vectores y solo falta normalizarlos. Un cálculo muestra $\vec{x}_1^t B \vec{x}_1 = 4$ y $\vec{x}_2^t B \vec{x}_2 = 12$, así que hay que dividir el primero por 2 y el segundo por $\sqrt{12} = 6/\sqrt{3}$. Así se obtiene finalmente la matriz de cambio

$$(26) \quad C = \begin{pmatrix} 1 & \sqrt{3}/3 \\ 1/2 & -\sqrt{3}/6 \end{pmatrix} \quad \text{que satisface} \quad C^t A C = \begin{pmatrix} -2 & 0 \\ 0 & 2/3 \end{pmatrix} \quad \text{y} \quad C^t B C = I.$$

Veamos ahora un caso con soluciones dobles.

Ejemplo. Tomamos las formas cuadráticas en \mathbb{R}^3 cuyas matrices son

$$(27) \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 3 \end{pmatrix}.$$

Se tiene $|A - \lambda B| = -\lambda^3 + 3\lambda + 2$ y por tanto $\lambda_1 = \lambda_2 = -1$ y $\lambda_3 = 2$. El núcleo de $A - \lambda_1 B$ tiene ahora consecuentemente dimensión dos y por tanto la libertad para elegir una base no se reduce a multiplicar por constantes. Tomemos por ejemplo $\vec{x}_1 = (1, -2, 1)^t$ y $\vec{x}_2 = (1, -1, 0)^t$. En el núcleo de $A - \lambda_3 B$

tomamos $\vec{x}_3 = (1, -1, 1)^t$, la única posibilidad salvo constantes. Por el proceso de Gram-Schmidt, los vectores \vec{x}_1 y \vec{x}_2 pasaran a ser ortogonales con el producto escalar dado por B si sustituimos \vec{x}_2 por

$$(28) \quad \vec{x}_2 - \frac{\langle \vec{x}_1, \vec{x}_2 \rangle}{\langle \vec{x}_1, \vec{x}_1 \rangle} \vec{x}_1 = \vec{x}_2 - \frac{1}{1} \vec{x}_1 = (0, -1, 1)^t.$$

Unos cálculos prueban que los vectores hallados $(1, -2, 1)^t$, $(0, -1, 1)^t$ y $(1, -1, 1)^t$ cumplen $\vec{x}^t B \vec{x} = 1$ y entonces dan sin más modificaciones las columnas de la matriz de cambio buscada.

Los enlaces atómicos a veces se modelan en primera aproximación como resortes mecánicos y la manera en la que oscila una molécula conduce a un problema de diagonalización simultánea. Veamos un análisis cualitativo de una de las situaciones más sencillas.

Ejemplo. Consideremos una molécula lineal con tres átomos, digamos que los de los extremos tienen masa m y el del centro es mayor con masa αm , $\alpha > 1$. Interpretemos los dos enlaces como muelles cuya constante es en ciertas unidades $\frac{1}{2}m$. Si x_1 , x_2 y x_3 son las separaciones de los átomos de sus posiciones de equilibrio, entonces la energía potencial de ambos muelles será $\frac{1}{2}m(x_1 - x_2)^2$ y $\frac{1}{2}m(x_2 - x_3)^2$. En definitiva, las energías potencial y cinética del sistema vienen dadas por

$$(29) \quad Q_p(\vec{x}) = \frac{1}{2}m\vec{x}^t A \vec{x} \quad \text{y} \quad Q_k(\vec{x}) = \frac{1}{2}m\vec{x}^t B \vec{x} \quad \text{con} \quad A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matriz A es la que corresponde a desarrollar $(x_1 - x_2)^2 + (x_2 - x_3)^2$. Resolviendo $|A - \lambda B| = 0$ se tiene $\lambda_1 = 0$, $\lambda_2 = 1$ y $\lambda_3 = 1 + 2\alpha^{-1}$. Estos valores en sí no tienen mucho significado porque las unidades son artificiales. Calculando vectores en los núcleos obtenemos $\vec{v}_1 = (1, 1, 1)^t$, $\vec{v}_2 = (1, 0, -1)^t$, $\vec{v}_3 = (1, -2\alpha^{-1}, 1)^t$. Esto da una información cualitativa interesante si recordamos que en esta base las oscilaciones son $(a_j \cos(t\sqrt{\lambda_j}) + b_j \sin(t\sqrt{\lambda_j}))\vec{v}_j$. Realmente el primer vector no corresponde a una oscilación porque $\lambda_1 = 0$ y la solución degenera a $(a_i + b_i t)\vec{v}_j$ que es una traslación de la molécula. El segundo corresponde a que el átomo central queda fijo mientras que los otros dos átomos se alejan y acercan al unísono (en la jerga se llama *breathing mode*, la molécula “respira” expandiéndose y contrayéndose). Finalmente el último vector indica una posibilidad de oscilación más complicada en la que el átomo grande se mueve hacia un lado y los pequeños hacia el otro siendo la razón de sus frecuencias al cuadrado $\lambda_3/\lambda_2 = 1 + 2\alpha^{-1}$.

Es posible diagonalizar simultáneamente si una de las formas es definida negativa con un proceso similar al descrito y un resultado caracteriza cuándo se puede llevar a cabo para formas no degeneradas. Aquí simplemente nos limitamos a ver un contraejemplo sencillo.

Ejemplo. Comprobemos que las formas cuadráticas en \mathbb{R}^2 dadas por $Q_1 = x^2 - y^2$, $Q_2 = 2xy$ no se pueden diagonalizar simultáneamente. Si pudiéramos escribir sus matrices como $C^t D_1 C$ y $C^t D_2 C$ con D_j matrices diagonales reales entonces

$$(30) \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (C^t D_1 C)^{-1} C^t D_2 C = C^{-1} D_1^{-1} D_2 C.$$

Esto implicaría que los elementos de la diagonal de $D_1^{-1} D_2$ son autovalores de la matriz de la izquierda, pero esta matriz no tiene autovalores reales.