Clasificación de Cuádricas

Gabriel Larotonda

Febrero 2025

1. Formas cuadráticas y bilineales, funciones homogéneas

A lo largo de este texto (salvo que se indique lo contrario) asumiremos que $\mathbb{K}$ es un cuerpo, de característica distinta de 2.

Definición 1.1. Sean V, W, Z $\mathbb{K}$-espacios vectoriales:

1. $\beta : V \times W \to Z$ es un *operador bilineal* si es lineal en cada variable por separado.
2. Para cada $v \in V$ fijo, denotamos $\beta_v : W \to Z$ al operador lineal $\beta_v(w) = \beta(v, w)$.
3. Cuando $V = W$, denotaremos $Q : V \to Z$ al *operador cuadrático asociada* a β, $Q(v) = \beta(v, v)$.
4. Si $\beta(v, w) = \beta(w, v)$ para todo $v, w \in V$, diremos que β es *simétrica*.
5. Si $\mathbb{K} = \mathbb{C}$, β es $\mathbb{C}$-lineal en su primer variable y $\beta(v, w) = \overline{\beta(w, v)}$ para todo $v, w \in V$, diremos que β es un operador *sesquilineal*.

Observemos que cuando cuando $V = W$, el operador cuadrático Q es una función homogénea de grado 2, esto es $Q(\lambda x) = \lambda^2 Q(x)$ para todo $x \in X$ y todo $\lambda \in \mathbb{K}$.

Sin embargo, en el caso de un operador sesquilineal, obtenemos que $Q(\lambda v) = \beta(\lambda v, \lambda v) = \lambda\bar{\lambda}Q(v) = |\lambda|^2 Q(v)$.

Observación 1.2. Cuando $Z = \mathbb{K}$, es usual denominar a los objetos que llegan a él *formas*. Así $\varphi \in X^*$ (el espacio dual de X) es una *forma lineal* en X, mientras que $\beta : X \times Y \to \mathbb{K}$ es una *forma bilineal* y $Q : X \to \mathbb{K}$ es una *forma cuadrática*.

Diremos que una forma bilineal sobre $\mathbb{R}$ (o sesqui-lineal) β es *semi-definida positiva* si $\beta(v, v) = Q(v) \geq 0$ para todo $v \in V$, y *definida positiva* si $\beta(v, v) = Q(v) = 0$ sólo cuando $v = 0$.
El *núcleo* de β es el subespacio

$$\ker\beta = \{v \in V : \beta(v, w) = 0 \, \forall w \in V\},$$

o equivalentemente el conjunto de todos los $v \in V$ tales que $\beta_v = 0$ como forma lineal en V.

1.1. Polarización

Cuando Q es el operador cuadrático dado por un operador bilineal $\beta : V \times V \to Z$, las *fórmulas de polarización* están dadas por

$$\begin{aligned} \beta_0(v, w) &= 1/2\{Q(v + w) - Q(v) - Q(w)\} \\ &= 1/4\{Q(v + w) - Q(v - w)\} \\ &= 1/2\{Q(v) + Q(w) - Q(v - w)\}, \end{aligned} \tag{1}$$

y definen *otro* operador bilineal $\beta_0 : V \times V \to Z$. Es evidente que β_0 es simétrica, y dejamos verificar al lector que

$$2\beta_0(v, w) = \beta(v, w) + \beta(w, v),$$

luego $\beta_0 = \beta$ si y sólo si β era originalmente simétrica.
Notemos que en el caso de que $\beta : V \times V \to \mathbb{R}$ es un producto interno sobre un $\mathbb{R}$-espacio vectorial,

$$\beta(v, w) = \langle v, w \rangle, \quad Q(v) = \beta(v, v) = \|v\|^2$$

de las identidades de polarización también obtenemos

$$2\|v\|^2 + 2\|w\|^2 = \|v + w\|^2 + \|v - w\|^2, \tag{2}$$

llamada *ley del paralelogramo*, porque relaciona las longitudes de los lados de un paralelogramo con las longitudes de sus diagonales.

Polarización en $\mathbb{C}$-espacios vectoriales

Observemos que si V es un espacio vectorial sobre los números complejos $\mathbb{C}$, la definición dada arriba de forma sesqui-lineal es especial ya que no se trata de una forma bilineal (sobre $\mathbb{C}$). De hecho resulta de la definición que β es $\mathbb{C}$ lineal en la primer variable, y $\mathbb{C}$ lineal conjugada en la segunda variable. Luego la identidad de polarización dada más arriba, no recupera β a partir de su forma cuadrática (sólo recupera la parte real de β).

Sin embargo, puede el lector verificar que la identidad de polarización

$$4\beta(v,w) = Q(v+w) - Q(v-w) - iQ(v-iw) + iQ(v+iw) \tag{3}$$

recupera la forma sesqui-lineal β a partir del operador cuadrático Q. Esta cuenta muestra, además, que toda forma sesqui-lineal se descompone como $\beta = \beta_0 + i\beta_1$ donde β_0, β_1 son formas bilineales simétricas reales.

1.2. De operadores cuadráticos a bilineales

Ahora abordaremos el problema recíproco, que puede pensarse como ¿cuándo una función homogénea de grado 2, denotada $Q : V \to Z$, proviene de un operador bilineal simétrico? O bien en particular ¿Cuándo una norma proviene de un producto interno?

Observemos primero que en general **esto no es cierto**: no toda función homogénea de grado 2 proviene de una forma bilineal. Es decir, dada Q función homogénea de grado 2, la identidad de polarización (1) no necesariamente define un operador bilineal β_0.

Consideremos $f : S^1 \to \mathbb{R}$ ($S^1 \subset \mathbb{R}^2$ es la circunferencia unitaria en el plano) una función simétrica: $f(-v) = f(v)$. Podemos definir una forma homogénea de grado 2 de la siguiente manera

$$Q(v) = \begin{cases} \|v\|^2 f\left(\frac{v}{\|v\|}\right) & v \neq 0 \\ 0 & v = 0. \end{cases}$$

En efecto, es evidente que $Q(0v) = Q(0) = 0 = 0^2 Q(v)$ para todo $v \in \mathbb{R}^2$, mietras que para $r > 0$, tenemos

$$Q(rv) = \|rv\|^2 f\left(\frac{rv}{\|rv\|}\right) = r^2\|v\|^2 f\left(\frac{v}{\|v\|}\right) = r^2 Q(v).$$

Similarmente

$$\begin{aligned} Q(-rv) =& \|-rv\|^2 f\left(\frac{-rv}{\|-rv\|}\right) = r^2\|v\|^2 f\left(-\frac{v}{\|v\|}\right) \\ =& r^2 Q(v) = (-r)^2 Q(v) \end{aligned}$$

usando la paridad de la función ($f(-w) = f(w)$).

1.2.1. Diferenciabilidad

También es evidente que si f es continua, será acotada y luego $Q : \mathbb{R}^2 \to \mathbb{R}$ será continua. Consideremos $f(x,y) = x^{2j}$ para $(x,y) \in \mathbb{R}^2$, $j \in \mathbb{N}_0$, ques es simétrica y continua. Obtenemos

$$Q(x,y) = \begin{cases} (x^2+y^2)\frac{x^{2j}}{(x^2+y^2)^j} = x^{2j}(x^2+y^2)^{1-j} & (x,y) \neq 0 \\ 0 & (x,y) = 0. \end{cases}$$

Podemos observar que cuando $j = 0$, f es constante y Q es la norma euclídea al cuadrado (que claramente proviene de una forma bilineal, la del producto interno usual de $\mathbb{R}^2$). Mientras que cuando $j = 1$, nos queda $Q(x,y) = x^2$.

Usando la fórmula de polarización (1) definimos $\beta : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ una función simétrica. Es inmediato verificar que $\beta(2v,w) \neq 2\beta(v,w)$, salvo los casos ya mencionados ($j = 0, j = 1$). En el segundo caso es porque se obtiene (polarizando $Q(x,y) = x^2$)

$$\beta(x,y;x',y') = xx'$$

que es en efecto una forma bilineal sobre $\mathbb{R}^2$. Notamos que tanto esta β como la dada por el producto interno usual ($j = 0$) son claramente funciones de clase C^∞ en $\mathbb{R}^2$.

En general, si f es de clase C^r, resultará lo mismo para Q en $\mathbb{R}^2 \setminus \{(0,0)\}$. Para $j \geq 2$, un cálculo directo sobre $Q(x,y) = x^{2j}(x^2+y^2)^{1-j}$ nos dice que $\frac{\partial Q}{\partial x}(0,0) = \frac{\partial Q}{\partial y}(0,0) = 0$ y que de hecho resulta Q diferenciable en $v = (0,0)$. Esto muestra que *la diferenciabilidad de Q no garantiza que provenga de una forma bilineal*.

Este ejemplo (o colección de ejemplos) que acabamos de presentar, nos muestra la relevancia del siguiente criterio, donde denotaremos D_wQ a la diferencial primera y D^2_wQ a la diferencial segunda de Q (evaluadas en $w \in V$):

Lema 1.3. *Sea $Q : V \to Z$ una función homogénea de grado 2. Si Q es de clase C^2 en un entorno de $v = 0$, entonces*

$$Q(v) = 1/2\, D^2_0 Q(v,v) \quad \forall v \in V.$$

En particular Q es un operador cuadrático proveniente del operador bilineal simétrico

$$\beta(v,w) = 1/2\, D^2_0 Q(v,w).$$

Demostración. La hipótesis dice que si $HQ = D\nabla Q$ es la matriz Hessiana de Q, esta resulta una matriz simétrica (ya que las derivadas cruzadas son iguales). Luego $\beta(v,w) = D^2_0 Q(v,w)$ es un operador bilineal simétrico. Veamos que coincide con $2Q$ en la diagonal. Para ello escribimos $Q(tv) = t^2 Q(v)$ para $t \in \mathbb{K}$, $v \in V$ y usando que la composición de funciones C^2 sigue siendo C^2, derivamos dos veces respecto del parámetro t (con v fijo, esto es válido para t suficientemente pequeño). Tenemos

$$D_{tv}Q(v) = 2tQ(v) \quad y \quad D^2_{tv}Q(v,v) = 2Q(v).$$

Evaluando en $t = 0$ tenemos la conclusión del lema. □

Observación 1.4. Cuando Q es de clase C^2 en un entorno de $v = 0$, el lema previo nos dice que $Q(v) = \beta(v,v)$ para β bilineal. Tomando $v_t = v(t) = v + tw$ y derivando esta expresión respecto de t, obtenemos

$$D_{v_t}Q(\dot{v}_t) = \beta(\dot{v}_t, v_t) + \beta(v_t, \dot{v}_t) = 2\beta(v_t, \dot{v}_t).$$

Evaluando en $t = 0$ se deduce que $D_vQ(w) = 2\beta(v,w)$. Luego $D_vQ = 2\beta(v,\cdot) = 2\beta_v$. Volviendo a reemplazar v por $v_t = v + tz$, obtenemos ahora derivando $D_vQ(w) = 2\beta(v,w)$ respecto de t (y evaluando en $t = 0$), que $D^2_v(z,w) = (D^2_vQ)(z)(w) = 2\beta(z,w)$, o equivalentemente $D^2_v = 2\beta$ para todo $v \in V$. Esto nos dice que la diferencial segunda de Q es constante (y de hecho igual a $D^2_0 Q$, como probamos en el lema). Se sigue que $D^3 Q \equiv 0$ y lo mismo para cualquier diferencial superior. Resumiendo, para todo $v \in V$, tenemos

$$D_vQ = 2\beta_v, \quad D^2_v = 2\beta, \quad D^k_v Q \equiv 0 \; \forall k \geq 3$$

y en particular Q es una función de clase C^∞ en V. Asismismo, como D^2Q es constante, es claro que Q coincide con su polinomio de Taylor de grado 2 centrado en cualquier $v \in V$:

$$\begin{aligned} Q(x) &= Q(v) + D_vQ(x-v) + 1/2\, D^2_v Q(x-v, x-v) \\ &= Q(v) + 2\beta_v(x-v) + Q(x-v), \end{aligned}$$

identidad que también es inmediata a partir de $Q(x-v) = \beta(x-v, x-v)$ y la bilinealidad de β.

1.2.2. Ley del (semi)-paralelogramo

Un resultado en la misma dirección pero sin recurrir a nociones de diferenciabilidad, se basa en la ley del paralelogramo. Como vimos, si una norma proviene de un producto interno, entonces cumple dicha ley (2). El siguiente teorema es su recíproco:

Teorema 1.5 (Teorema de Jordan-von Neumann)**.** *Si* $\|\cdot\| : V \to \mathbb{R}$ *es una norma que obedece la ley del semi-paralelogramo*

$$2\|v\|^2 + 2\|w\|^2 \leq \|v+w\|^2 + \|v-w\|^2 \quad \forall v, w \in V,$$

entonces la polarización β *de la norma da un producto interno. En consecuencia, una norma proviene de un producto interno si y sólo si obedece la ley del semi-paralelogramo.*

Demostración. En primer lugar, observemos que reemplazando v por $1/2(v+w)$ y w por $1/2(v-w)$, la desigualdad es equivalente a la igualdad. Esto es, podemos suponer que vale la ley del paralelogramo. Definimos el producto interno $\beta(v,w) = \langle v,w\rangle$ usando la fórmula de polarización (1). Claramente es simétrico. Tenemos que probar que se trata de una forma bilineal. Ahora notamos que, por la ley del paralelogramo aplicada a $v+z, w$ (y a $w+z, v$), se tiene

$$\begin{aligned}\|v+w+z\|^2 &= 2\|v+z\|^2 + 2\|w\|^2 - \|v-w+z\|^2 \\ &= 2\|w+z\|^2 + 2\|v\|^2 - \|w-v+z\|^2.\end{aligned}$$

Luego sumando y dividiendo por 2

$$\begin{aligned}\|v+w+z\|^2 =& \|v\|^2 + \|w\|^2 + \|v+z\|^2 + \|w+z\|^2 \\ & - \frac{1}{2}\|v-w+z\|^2 - \frac{1}{2}\|w-v+z\|^2.\end{aligned}$$

La misma identidad vale para $-z$. Entonces restando estas dos experesiones (la de z y la de $-z$), obtenemos $\beta(v+w,z)$ de acuerdo a la identidad de polarización (1) que usamos para definir el producto interno. Esto es

$$\begin{aligned}\langle v+w, z\rangle &= \frac{1}{4}\left(\|v+w+z\|^2 - \|v+w-z\|^2\right) \\ &= \frac{1}{4}\left(\|v+z\|^2 - \|v-z\|^2\right) + \frac{1}{4}\left(\|w+z\|^2 - \|w-z\|^2\right) \\ &= \langle v,z\rangle + \langle w,z\rangle\end{aligned}$$

donde en la primer igualdad usamos la hipótesis de la regla del paralelogramo, y en la última igualdad usamos polarización para obtener $\beta(v,z)$ y $\beta(w,z)$. Esto prueba que β es aditiva en la primer variable, y por simetría lo es en la segunda. Nos falta ver que β saca escalares reales (y por simetría alcanza con verlo en la primer variable). De la desigualdad triangular inversa para la norma,

$$|\,\|v\| - \|w\|\,| \leq \|v-w\|$$

es inmediato que la norma es una función continua, luego β es continua. Entonces bastará probar que β saca escalares racionales, es decir $\lambda = p/q \in \mathbb{Q}$. Como $\beta(0,w) = 0$, y de la definición es inmediato que

$$4\langle -v, w\rangle = \|-v+w\|^2 - \|-v-w\|^2 = \|w-v\|^2 - \|v+w\|^2 = -4\langle v,w\rangle,$$

podemos suponer que $\lambda > 0$, esto es $p,q \in \mathbb{N}$. Como ya probamos que β es aditiva en la primer variable, en realidad ya lo probamos si $q = 1$. Ahora bien si $v' = 1/q\, v$, tenemos

$$q\langle \lambda v, w\rangle = q\langle pv', w\rangle = pq\langle v', w\rangle = p\langle qv', w\rangle = p\langle v,w\rangle.$$

Dividiendo por q obtenemos lo deseado, $\langle \lambda v, w\rangle = \lambda\langle v,w\rangle$. □

El caso de productos internos complejos (formas sesquilineales)

En el caso de que el espacio vectorial V sea un $\mathbb{C}$ espacio vectorial, dad una norma sobre V, podemos definir

$$\begin{aligned} 4\langle v,w\rangle &= \|v+w\|^2 - \|v-w\|^2 + i[\|v+iw\|^2 - \|v-iw\|^2] \\ &= \beta_0(v,w) + i\beta_1(v,w) \end{aligned}$$

que no es otra cosa que la identidad de polarización compleja (3). Es inmediato de la definición que $\langle w,v\rangle = \overline{\langle v,w\rangle}$, y que $\langle iv,w\rangle = i\langle v,w\rangle$.

Si suponemos que la norma obedece la propiedad del semi - paralelogramo, sabemos que obedece la ley del paralelogramo. No es difícil ver que entonces β_0 y β_1 son ambas $\mathbb{R}$ bilineales, por el Teorema de Jordan-Von Neumann. Luego el producto interno definido arriba resulta un producto interno complejo, esto es, una forma sesqui-lineal.

Corolario 1.6. *Sobre un $\mathbb{C}$ espacio vectorial, una norma obedece la ley del semi- paralelogramo si y sólo si proviene de un producto interno complejo. Es decir, si la norma proviene de una forma sesqui-lineal definida positiva.*

1.3. Cono asociado a una forma cuadrática

Si $Q : V \to \mathbb{K}$ es una forma cuadrática, el cono asociado se define como

$$C_Q = \{v \in V : Q(v) = 0\}.$$

Nos interesan dos propiedades fundamentales de los conos que utilizaremos para la clasificación de cuádricas. Estas propiedades dicen que fuera de un cono hay infinitos hiperplanos. Todas las demás propiedades de C_Q pueden deducirse luego, utilizando la clasificación de cuádricas.

Lema 1.7. *Si $Q \neq 0$ y $H \subset C_Q$ es un hiperplano dado por una funcional $\varphi \in V^*$, entonces existe $\psi \in V^*$ tal que $Q = \varphi\psi$. Luego $C_Q = H \cup \tilde{H}$ donde $\tilde{H} = \ker\psi$ puede ser igual a H. En particular C_Q contiene a lo sumo dos hiperplanos.*

Demostración. Ejercicio. □

Lema 1.8. *Si $Q \neq 0$ y S_1, S_2 son subespacios propios de V, entonces $S_1 \cup S_2 \cup C_Q \neq V$.*

Demostración. La demostración es por inducción en $d = \dim(V)$. Si $d = 1$, debe ser $S_1 = S_2 = \{0\}$ y entonces $S_1 \cup S_2 \cup C_Q = C_Q \neq V$ pues $Q \neq 0$. Supongamos que hemos probado el resultado para $d \geq 1$, sean $H_1, H_2 \subset V$ hiperplanos tales que $H_i \neq S_1$ y C_Q no contiene H_i ($i = 1,2$), que existen por el lema previo. Esto es lo mismo que decir que $Q|_{H_i} \neq 0$. Entonces si fuese $V = S_1 \cup S_2 \cup C_Q$, tendríamos

$$H_1 = (H_1 \cap S_1) \cup (H_1 \cap S_2) \cup (H_1 \cap C_Q) = \tilde{S_1} \cup \tilde{S_2} \cup C_{\tilde{Q}}$$

donde $\tilde{Q} = Q|_{H_1}$. Notemos que $\tilde{S_1} \neq H_1$ y que $\tilde{Q} \neq 0$. Como $\dim(H_1) = \dim(V) - 1$, por la hipótesis inductiva la única posibilidad es que $\tilde{S_2} = H_1$, es decir $H_1 \cap S_2 = H_1$. Luego $H_1 \subset S_2$ pero como H_1 es un hiperplano debe ser $S_2 = H_1$. Con un razonamiento análogo, descomponiendo H_2, debe ser también $S_2 = H_2$. Pero entonces $H_1 = H_2$, absurdo. □

2. Representación de formas cuadráticas fijando una base

Dada una forma bilineal β en un $\mathbb{K}$-e.v. V, queremos dar una representación de β en términos de una transformación lineal $T : V \to V$. Para ello notamos que, fijar una base $B = \{v_1, \dots v_n\}$ de V es equivalente a elegir un producto interno en V.

Recordemos que el método de Gram-Schmidt permite obtener una base ortonormal para cualquier producto interno, y recíprocamente, dada una base ortonormal podemos definir un producto interno decretando que estos vectores formen una base ortonormal. Vamos a enunciar esto de manera más precisa (la verificación es trivial y queda a cargo del lector).

Lema 2.1. *Dado un producto interno en V, existe una base ortonormal $E = \{e_1, \ldots, e_n\}$ para éste, y toda base ortonormal para este producto interno se obtiene a partir de la misma mediante una transformación ortogonal: esto es una transformación lineal $U : V \to V$ que es un isomorfismo, tal que $\|Uv\| = \|v\|$ para todo $v \in V$, y tal que $Ue_i = v_i$.*
Recíprocamente fijada una base arbitraria E de V, si escribimos, para $v, w \in V$ sus coordenadas en esta base,

$$v = \sum_{i=1}^{n} x_i e_i, \quad w = \sum_{i=1}^{n} y_i e_i, \tag{4}$$

definimos

$$\langle v, w \rangle = \sum_{i=1}^{n} x_i y_i,$$

que es un producto interno que hace de E una base ortonormal para el mismo. En el caso de un espacio vectorial sobre $\mathbb{C}$, podemos obtener un producto interno complejo definiendo

$$\langle v, w \rangle = \sum_{i=1}^{n} x_i \bar{y}_i.$$

Una vez fijado un producto interno en V (equivalentemente, fijada una base de V), dado un subespacio $S \subset V$, denotaremos $S^\perp \subset V$ al *complemento ortogonal* de S,

$$S^\perp = \{v \in V : \langle v, s \rangle = 0 \quad \forall s \in S\}.$$

2.1. Transformaciones lineales autoadjuntas

Recordemos que una transformación lineal $T : V \to V$ es *autoadjunta* si

$$\langle Tv, w \rangle = \langle v, Tw \rangle \quad \forall v, w \in V.$$

Es inmediato de aquí que el rango (la imagen) y el núcleo de T son ortogonales, y entonces

$$V = \ker T \oplus \mathrm{Ran}(T)$$

donde $\ker T = \mathrm{Ran}(T)^\perp$.
En general, si T es una transformación lineal, su adjunta se define como la única transformación lineal T^* tal que

$$\langle Tv, w \rangle = \langle v, T^* w \rangle \quad \forall v, w \in V.$$

Probaremos ahora dos teoremas de representación debidos a F. Riesz. Diremos que una aplicación aditiva R entre $\mathbb{C}$ espacios vectoriales es *anti-lineal* si $R(\lambda z) = \overline{\lambda} R(z)$ para todo vector z y todo $\lambda \in \mathbb{C}$.

Observación 2.2. Notemos que dado un vector $y \in V$, este induce una forma lineal $\varphi_y \in V^*$, la dada por $\varphi_y(v) = \langle v, y \rangle$. La aplicación $y \to \varphi_y$ es claramente lineal e inyectiva (anti-lineal cuando $\mathbb{K} = \mathbb{C}$). Si definimos la norma de una funcional como

$$\|\varphi\| = \max\{|\varphi(v)| : v \in V, \ \|v\| = 1\},$$

es inmediato (por la desigualdad de Cauchy-Schwarz) que esta aplicación es una isometría entre V y V^* (preserva las normas).

El primer teorema de Riesz dice que se trata de una biyección:

Teorema 2.3 (Teorema de representación de Riesz para formas lineales)**.** *Sea $\varphi \in V^*$, entonces existe un único $y_\varphi \in V$ tal que $\varphi = \langle \cdot, y_\varphi \rangle$. La aplicación $R(\varphi) = y_\varphi$ se lo denomina mapa de Riesz; es un isomorfismo isométrico cuando $\mathbb{K} = \mathbb{R}$, mientras que es un anti-isomorfismo isométrico cuando $\mathbb{K} = \mathbb{C}$.*

Demostración. Sea $S = \ker \varphi$, si $S = V$ es porque $\varphi = 0$, tomamos $y_\varphi = 0$. En caso contrario, S es un subespacio propio (un hiperplano). Tomamos $y_0 \in S^\perp$ de norma unitaria y definimos

$$y_\varphi = \overline{\varphi(y_0)} y_0$$

(por supuesto que la conjugación no es necesaria cuando $\mathbb{K} = \mathbb{R}$). Descomponemos $x \in V$ como $x = s + \lambda y_\varphi$, notamos que $\varphi(x) = 0 + \lambda \varphi(y_\varphi) = \lambda |\varphi(y_0)|^2$ pues φ es lineal. Por otro lado,

$$\langle x, y_\varphi \rangle = 0 + \lambda \langle y_\varphi, y_\varphi \rangle = \lambda \|y_\varphi\|^2 = \lambda |\varphi(y_0)|^2.$$

Esto prueba que $\varphi = \langle \cdot, y_\varphi \rangle$ y por la observación previa, tienen la misma norma. La unicidad queda a cargo del lector. $\square$

El segundo teorema de Riesz dice que toda forma bilineal simétrica (o sesqui-lineal) en V se puede representar mediante una transformación lineal $T : V \to V$. Enunciamos el teorema para $\mathbb{K} = \mathbb{C}$, el análogo para $\mathbb{R}$ es inmediato omitiendo los conjugados.

Teorema 2.4 (Teorema de representación de Riesz para formas bilineales)**.** *Sea $\beta : V \times V \to \mathbb{C}$ sesquilineal. Existe entonces una única transformación lineal autoadjunta $T : V \to V$ tal que*

$$\beta(x, y) = \langle x, Ty \rangle = \langle Tx, y \rangle \qquad \forall x, y \in V.$$

Demostración. Para cada $y \in V$, sea $\varphi_y(x) = \beta(x, y)$. Luego existe un único $z = T(y) \in V$ tal que

$$\beta(x, y) = \varphi(x) = \langle x, z \rangle = \langle x, Ty \rangle \qquad \forall x \in V.$$

Veamos primero que T es lineal. Para ello, calculamos, para cada $\lambda \in \mathbb{C}$

$$\langle x, T(\lambda y) \rangle = \beta(x, \lambda y) = \overline{\lambda} \beta(x, y) = \overline{\lambda} \langle x, Ty \rangle = \langle x, \lambda Ty \rangle.$$

Como esto vale para todo $x \in V$, se tiene $T(\lambda y) = \lambda Ty$ para todo $y \in V, \lambda \in \mathbb{C}$. Con un argumento similar, y usando que el mapa de Riesz es aditivo, se obtiene que T es aditiva. Como β es sesquilineal,

$$\overline{\langle Tx, y \rangle} = \langle x, Ty \rangle = \beta(x, y) = \overline{\beta(y, x)} = \overline{\langle y, Tx \rangle},$$

de donde resulta inmediato que T es autoadjunta. $\square$

Ahora probaremos el resultado fundamental sobre transformaciones lineales autoadjuntas:

Teorema 2.5. *Si $T : V \to V$ es autoadjunta, existe una base ortonormal $E = \{e_1, \ldots, e_n\}$ de V de autovectores de T, todos los autovalores de T son reales, y los autoespacios correspondientes a autovalores distintos son ortogonales, e invariantes para T.*

Demostración. Como $T^* = T$, entonces si λ es autovalor, y v es autovector, debe ser

$$\lambda \|v\|^2 = \langle Tv, v \rangle = \langle v, Tv \rangle = \overline{\langle Tv, v \rangle} = \overline{\lambda} \|v\|^2$$

lo que prueba que todo autovalor es real.

Notamos ahora que si v es autovector, entonces $S = \{v\}^\perp$ es invariante para T: en efecto, si tomamos $w \perp v$, entonces

$$\langle Tw, v \rangle = \langle w, Tv \rangle = \lambda \langle w, v \rangle = 0.$$

Consideramos la restricción de T al subespacio $\{v\}^{\perp}$, donde sigue siendo autoadjunta. Procedemos inductivamente hasta construir una base ortonormal de V de autovectores de T. Todos los autovalores son reales, y si $\lambda \neq \mu$ son autovalores de T, entonces, tomando v, w autovectores respectivos,

$$\lambda\langle v, w\rangle = \langle Tv, w\rangle = \langle v, Tw\rangle = \mu\langle v, w\rangle,$$

lo cuál solo es posible si $v \perp w$. □

Definición 2.6 (Descomposición espectral de una t.l. autoadjunta)**.** Dado λ autovalor de T autoadjunta, denotaremos $E_\lambda^T \subset V$ al *autoespacio asociado de autovectores*, esto es $E_\lambda^T = \{v \in V : Tv = \lambda v\}$. Notemos que se trata de un subespacio invariante por T, y que los E_λ^T son todos ortogonales entre si (para cada autovalor de T). Denotaremos P_λ^T al único proyector ortogonal con rango E_λ^T, lo llamamos *proyector espectral de* T.

Si λ_i son los autovalores *distintos* de T (incluyendo $\lambda = 0$ posiblemente), y denotamos $E_i = E_{\lambda_i}^T$, $P_i = P_{\lambda_i}^T$, afirmamos que

$$T = \sum_i \lambda_i P_i.$$

En efecto, dado $v \in V$, este tiene una única escritura $v = \sum v_i$ con $v_i = P_i v \in E_i$. Luego

$$Tv = \sum_i Tv_i = \sum_i \lambda_i v_i = \sum_i \lambda_i P_i v = (\sum_i \lambda_i P_i)v.$$

Notar que $1 \leq i \leq k$ con $k \leq n = \dim(V)$, ya que recorre los autovalores *distintos*.

Si denotamos $\sigma(T) = \{t \in \mathbb{R} : T - t1 \text{ no es inversible}\}$ al conjunto de todos los autovalores de T (denominado *espectro de* T), podemos denotar P_t al proyector asociado a cada autoespacio E_t^T, y entonces tenemos la reescritura

$$T = \sum_i \lambda_i P_i = \sum_{t\in\sigma(t)} t\, P_t \tag{5}$$

conocida como *descomposición espectral de* T. Esta es una versión (para espacios vectoriales de dimensión finita con producto interno) del llamado *teorema espectral para operadores autoadjuntos*.

Definición 2.7 (Cálculo funcional de una t.l. autoadjunta)**.** Como autoespacios correspondientes a autovalores distintos están en suma directa (son ortogonales), se tiene $P_iP_j = 0$ si $i \neq j$. Luego si $n \in \mathbb{N}$, es inmediato que $T^n = \sum_t t^n P_t$, y que si q es cualquier polinomio $q(x) = a_n x^n + \cdots + a_1 x + a_0$ entonces

$$q(T) = a_n T^n + \cdots + a_1 T + a_0 1 = \sum_t q(t)\, P_t.$$

Dada cualquier función $f : \mathbb{R} \to \mathbb{R}$, podemos *definir*

$$f(T) = \sum_{t\in\sigma(T)} f(t)\, P_t,$$

esta transformación lineal se denomina *cálculo funcional* de T evaluado en f. Si q_n es una sucesión de polinomios que tiende puntualmente a la función f, invitamos al lector a demostrar que $q_n(T) \to f(T)$ en cualquier norma que uno quiera ponerle al espacio de las transformaciones lineales.

Si P_0 denota el proyector ortogonal al núcleo de T, entonces tenemos las fórmula

$$V = \oplus_{t\in\sigma(T)} E_t^T$$

(donde cada subespacio es invariante para T, y son todos ortogonales entre sí), que resulta equivalente a la descomposición de la t.l. identidad $1 : V \to V$ como

$$1 = \sum_{t\in\sigma(T)} P_t,$$

y los proyectores verifican $P_t P_s = P_s P_t = 0$ si $s \neq t$.

Observación 2.8. Cabe notar aquí, que si U es transformación ortogonal y $E \subset V$ es un subespacio (y P_E su proyector ortogonal), entonces si $E' = U(E)$ se tiene

$$P_{E'} = UP_E U^*.$$

En efecto, $P_{E'}^2 = UP_E U^* UP_E U^* = UP_E P_E U^* = UP_E U^* = P_{E'}$, $P_{E'}^* = UP_E^* U^* = UP_E U^* = P_{E'}$ luego $P_{E'}$ es un proyector ortogonal. Por otro lado si $Uv \in U(E)$, entonces $P_{E'}(Uv) = UP_E U^* Uv = UP_E v = Uv$ pues $v \in E$, luego $P_{E'}(U(E)) = U(E)$. Mientras que si $z = Uy \in U(E)^\perp = U(E^\perp)$ entonces $P_{E'}z = UP_E U^* Uy = UP_E y = U0 = 0$, luego $\ker(P_{E'}) = U(E)^\perp$.

Observación 2.9. Si $B = \{v_1, \ldots, v_d\}$ es una base de V, podemos denotar $p_i = v_i \otimes v_i$ al proyector ortogonal con rango v_i, de manera tal que $\sum_i p_i = 1$, $p_i p_j = 0$ si $i \neq j$. Definimos una transformación lineal $Tv_i = \lambda_i v_i$, con $\lambda_i \in \mathbb{C}$, es decir $T = \sum_i \lambda_i p_i$. Resulta T autoadjunta: en efecto si $v = \sum_i x_i v_i$, $w = \sum_i y_i v_i$ entonces

$$\langle Tv, w \rangle = \sum_{i,j} x_i \overline{y_i} \lambda_i = \langle v, Tw \rangle.$$

Con la misma demostración, si $1 = \sum_i P_i$ es cualquier descomposición de la identidad $1 : V \to V$ en proyectores ortogonales disjuntos ($P_i P_j = 0$ si $i \neq j$), resultará $T = \sum_i \lambda_i P_i$ una t.l. autoadjunta, que es inversible si y sólo si todos los λ_i son no nulos.

Corolario 2.10. *Si T, R son t.l. autoadjuntas y conmutan, entonces se diagonalizan simultáneamente, es decir tienen una base ortonormal común de autovectores. Es más, los autoespacios de una son invariantes para el otro.*

Demostración. Dado λ autovalor de T, tenemos para cada $v \in E_\lambda^T$ (incluyendo $E_0^T = \ker T$)

$$TRv = RTv = R\lambda v = \lambda Rv,$$

lo que nos dice que $Rv \in E_\lambda^T$, o sea que E_λ^T es invariante también para R. Como R restringido a E_λ^T sigue siendo autoadjunto, tomamos una b.o.n. de E_λ^T pero conformada por autovectores de R (teorema anterior). Como son elementos de E_λ^T, resultan simultáneamente autovectores de T, de autovalor λ (notar que sin embargo, como autovectores de R, sus autovalores serán distintos en general). Podemos hacer esto para cada autoespacio de T, y $V = \oplus_\lambda E_\lambda^T$, obtenemos una base de autovectores de R que es también base de autovectores de T. □

Corolario 2.11. *Si T es autoadjunta, entonces R conmuta con T si y sólo si R conmuta con todos los proyectores espectrales de T.*

Demostración. Si R conmuta con todos los proyectores espectrales de T, entonces

$$RT = R \sum_{t \in \sigma(T)} tP_t = \sum_{t \in \sigma(T)} tRP_t = \sum_{t \in \sigma(T)} tP_t R = TR.$$

Recíprocamente, si R conmuta con T entonces $RT^n = T^n R$ para todo $n \in \mathbb{N}$, y con esto

$$\sum_t t^n RP_t = RT = TR = \sum_t t^n P_t R.$$

Dado $t_0 \in \sigma(T)$, sea f un polinomio que se anula en $\sigma(T) \setminus \{t_0\}$ y tal que $f(t_0) = 1$. Entonces

$$\sum_t f(t) RP_t = \sum_t f(t) P_t R,$$

luego $RP_{t_0} = P_{t_0} R$ como queríamos probar. □

2.2. Transformación autoadjunta asociada a una forma cuadrática

Si $Q : V \to \mathbb{R}$ es una forma cuadrática con forma bilineal asociada β, y fijamos una base de V, estaremos fijando un sistema de coordenadas y también un producto interno en V. Para este producto interno, existe una única transformación lineal autoadjunta $T : V \to V$ tal que

$$Q(v) = \langle Tv, v\rangle \quad \forall v \in V.$$

Dejamos al lector probar que $\ker T = \ker \beta$. Con esto, el núcleo de T no depende del producto interno elegido.

Al cambiar de base (cambiamos de coordenadas), la transformación T no será la misma (aunque tendrá el mismo núcleo). Sea $M : V \to V$ isomorfismo que intercambia estas bases, digamos $e_i = Mv_i$, $i = 1, \dots, n$, donde los e_i son los de la primer base y v_i los de la segunda. Luego si $v, w \in V$, con coordenadas como en (4), tendremos

$$v = \sum_i x_i e_i = \sum_i x_i M v_i = M \sum_i x_i v_i = Mv'.$$

Esto es si $v' = \sum_i x_i v_i$ entonces $v = Mv'$. Similarmente, $Mw = w'$. Por definición

$$\langle Mv', Mw'\rangle = \langle v, w\rangle = \sum_i x_i y_i = \langle v', w'\rangle_B$$

donde a la derecha tenemos el producto interno definido por la nueva base B. Como esta fórmula es genérica, se deduce

Lema 2.12. *Si E, B son dos bases de V y $M : V \to V$ es el isomorfismo $Mv_i = e_i$, entonces los productos internos están relacionados por la fórmula*

$$\langle v, w\rangle_B = \langle Mv, Mw\rangle = \langle M^*Mv, w\rangle \tag{6}$$

para todo $v, w \in V$.

Observación 2.13. Podríamos haber arribado a una identidad de este tipo usando el Teorema 2.4, ya que el producto interno en la base B es una forma bilineal, luego se tiene que representar por medio de una t.l. autoadjunta A, en términos del producto interno original (en este caso $A = M^*M$).

Observación 2.14. Si T^B indica la t.l. (autoadjunta para el nuevo producto interno) que determina únicamente la forma bilineal $\beta : V \times V \to \mathbb{K}$ asociada a Q, de la identidad (6) tenemos

$$\langle M^*MT^B v, w\rangle = \langle T^B v, w\rangle_B = \beta(v, w) = \langle Tv, w\rangle$$

para todo $v, w \in V$, luego $T = M^*MT^B$. Como M^*M es inversible, es inmediato que, como comentamos al comienzo de esta sección, T^B y T tienen el mismo núcleo.

2.3. Transformaciones (semidefinidas) positivas

Lo remarcable de transformación lineal $A = M^*M$ del Lema 2.12 es que no sólo es inversible (ya que M lo es) sino que todos sus autovalores son estrictamente positivos: si v es autovector de norma unitaria para el autovalor $\lambda \in \mathbb{R}$ de A, entonces

$$\lambda = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle M^*Mv, v\rangle = \langle Mv, Mv\rangle = \|Mv\|^2 \geq 0.$$

En este caso λ no puede ser nulo ya que A es inversible.

Definición 2.15. Dado un producto interno en V, diremos que una t.l. $A : V \to V$ es *positiva* (denotado $A \geq 0$) si A es autodajunta y todos los autovalores de A son no negativos. Diremos que A es *positiva inversible* (denotado $A > 0$) si A es autoadjunta y todos sus autovalores son estrictamente positivos (en particular es inversible).

Corolario 2.16. *Si $\langle\cdot,\cdot\rangle_0, \langle\cdot,\cdot\rangle_1$ son dos productos internos en V, existe una (única) t.l. $A : V \to V$ positiva inversible que los relaciona:*

$$\langle v,w\rangle_1 = \langle Av,w\rangle_0.$$

Definición 2.17 (Módulo)**.** Si T es autoadjunta y $T = \sum_{t\in\sigma(T)} tP_t$ es su descomposición espectral (Definición 2.7), entonces tomando $f(x) = |x|$, la matriz

$$|T| = \sum_{t\in\sigma(T)} |t|P_t,$$

se denomina *módulo* de T, y es inmediato que $|T|$ es positiva, y será inversible si y sólo si T lo era.

Lema 2.18 (Lema de la raíz cuadrada)**.** *Si V es un espacio con producto interno y $A \geq 0$, existe una única t.l. $R \geq 0$ tal que $A = R^2$. Se suele denotar $R = \sqrt{A} = A^{1/2}$ la "raíz cuadrada de A". Si $A = \sum_i a_iP_i = \sum_{a\in\sigma(a)} aP_a$ es la descomposición espectral de A (5), entonces $R = \sqrt{A} = \sum_i \sqrt{a_i}P_i = \sum_{a\in\sigma(a)} \sqrt{a}P_a$.*

Demostración. Como A es autoadjunta y $\lambda_i \geq 0$ para todos sus autovalores, si v_i es la base correspondiente de autovectores, definimos $Rv_i = \sqrt{\lambda_i}v_i$. Es inmediato que $R^2 = A$, y que $R \geq 0$. Veamos loa unicidad: si $A = T^2$ con $T \geq 0$, entonces $TA = TT^2 = T^2T = AT$ es decir T conmuta con A. Como ambas son autoadjuntas, existe una base $B = \{v_1,\dots,v_n\}$ de autovectores de A que también diagonaliza T (y que son base también de los autoespacios $E_a = E_a^A$, con $a \geq 0$ autovalor de A). Observemos que si $v \in E_a$, escribiendo $v = \sum_j \alpha_jv_j$ donde los v_j son algunos de los vectores de B, que están en E_a, tenemos $Tv = T\sum_j \alpha_jv_j = \sum_j \alpha_jt_jv_j$ donde $t_j \in \sigma(T)$. Pero aplicando T una vez más,

$$\sum_j \alpha_j av_j = av = Av = T^2v = \sum_j \alpha_jt_j^2v_j$$

y esto sólo es posible si $t_j = \sqrt{a}$ para todos estos j. Luego $Tv = \sqrt{a}\sum_j \alpha_jv_j = \sqrt{a}v$. Esto prueba que $TP_a = \sqrt{a}P_a$, y entonces

$$T = T1 = T\sum_{a\in\sigma(a)} P_a = \sum_{a\in\sigma(a)} \sqrt{a}P_a.$$

De aquí se deducen la unicidad de la raíz cuadrada y la última fórmula del enunciado del lema. □

Dejamos el siguiente problema para el lector:

Problema 2.19. Sea $A : V \to V$ t.l., V un espacio vectorial con producto interno. Probar

1. A es positiva $\Leftrightarrow$ $\langle Av,v\rangle \geq 0$ para todo $v \in V$ $\Leftrightarrow$ $\exists M$ tal que $A = M^*M$.
2. Probar que "$A \geq B$" si $A - B \geq 0$, es un orden parcial en el conjunto de las t.l.
3. Probar que si A es positiva e inversible, entonces A^{-1} también es positiva.
4. Probar que si A es positiva entonces MAM^* también lo es.

Definición 2.20 (Módulo y descomposición polar)**.** Dada T t.l. sabemos entonces que $T^*T \geq 0$ tiene una única raíz cuadrada positiva, que denotaremos $|T|$, esto es $|T| = \sqrt{T^*T}$ es la única t.l. M tal que $M \geq 0$ y $M^2 = T^*T$. Es claro que $|T|$ será inversible si y sólo si T lo era (notar que es la misma t.l. definida en 2.17 cuando $T^* = T$).

Si T es inversible, definimos la transformación inversible $U = T|T|^{-1}$. Tenemos

$$U^*U = |T|^{-1}T^*T|T|^{-1} = |T|^{-1}|T|^2|T|^{-1} = 1,$$

lo que prueba que U es una transformación *ortogonal*. La fórmula $T = U|T|$ se conoce como *descomposición polar* de T.

Definición 2.21. Dada una transformación lineal $T : V \to V$ definimos el rango $r = r(T)$ como la cantidad de autovalores no nulos, o equivalentemente, la dimensión de $\mathrm{Ran}(T) = \{Tv : v \in V\}$.

Cuando T es autoadjunta, todos sus autovalores son reales, y entonces definimos la *positividad* $p = p(T)$ como la cantidad de autovalores positivos, la *negatividad* $n = n(T)$ como la cantidad de autovalores negativos, y la *signatura* $s = s(T)$ como la diferencia entre ambos, $s = p - n$. Notar que dos cualesquiera de estos números determinan los otros dos, ya que $r = p + n$ luego $s + r = 2p$. Definimos el *índice* de T como el par $(p, r) \in \mathbb{N}_0 \times \mathbb{N}_0$. Notar que $\dim(\ker(T)) = n - r$.

Lema 2.22 (Descomposición de Hahn de una t.l. autoadjunta)**.** *Si T es autoadjunta, la descomposición $T = T_+ - T_-$ dada por*

$$T_+ = \sum_{t>0,\in\sigma(T)} t\,P_t, \qquad T_- = \sum_{t<0} (-t)P_t$$

verifica $T_+, T_- \geq 0$, $T_+T_- = T_-T_+ = 0$, $T_+ + T_- = |T|$.
En el subespacio invariante $Ran(T_+)$ se tiene $T > 0$, este subespacio es maximal respecto de la positividad de T, y su dimensión coincide con $p(T)$, el número de autovalores positivos de T. Mismas afirmaciones valen para el rango de T_-, que es ortogonal al rango de T_+, y además $Ran(T) = Ran(T_+) \oplus Ran(T_-)$.

Demostración. Que $T_+, T_- \geq 0$ es evidente porque son autoadjuntas y tienen autovalores positivos. Que los rangos de T_+, T_- son ortogonales (y entonces $T_+T_- = 0$) es inmediato de que autovectores correspondientes a autovalores distintos son ortogonales, y también es clara la última afirmación sobre el rango de T. Veamos que $T|_{\mathrm{Ran}(T_+)} > 0$; si dicho subespacio es nulo no hay nada que probar. Caso contrario, tomamos una base ortornomal que diagonaliza T y tomamos los primeros $v_1, \cdots, v_k$ que forman una base ortonormal de autovectores con todos los autovalores positivos de T, es evidente que

$$T = T_+ - T_- = \sum_{i=1}^{k} r_i v_i - \sum_{i=k+1}^{n} (-s_j) v_j$$

donde r_i son los autovalores negativos de T y s_j los negativos (si $k = n$ la segunda suma es nula y tanto T_- como su rango son nulos). Sea $0 \neq v \in \mathrm{Ran}(T_+)$, entonces $v = \sum_{i=1}^{k} x_i v_i$, luego $Tv_i = \sum_i x_i r_i v_i$ y con ello

$$\langle Tv, v\rangle = \sum_{i,j=1}^{k} r_i x_i x_j \langle v_i, v_j\rangle = \sum_{i=1}^{k} r_i x_i^2 > 0$$

pues algún x_i es no nulo. Supongamos ahora que $S \subset V$ es un subespacio invariante para T, entonces T se diagonaliza sobre S y si $T|_S$ es positivo, entonces todos los autovalores de T alli son estrictamente positivos, luego $\dim(S) \leq p(T)$ (porque sino habría más autovalores estrictamente positivos!). □

Veamos que los cambios de coordenadas no alteran estas cantidades, principio conocido como la *Ley de inercia de Sylvester*. Para ello probamos antes un lema que dice que transformaciones cercanas tienen autovalores cercanos, o en otros términos, que la función espectro $\sigma : End(V) \to 2^{\mathbb{C}}$ es semicontinua inferiormente.

Para ello, consideramos $(End(V), d)$ como espacio métrico, con la distancia dada por $d(T, R) = \|T - R\|$, y $\|\cdot\|$ es alguna norma en los endomorfismos, como por ejemplo, fijando una norma en V,

$$d(T, R) = \|T - R\|_\infty = \max\{\|Tv - Rv\| : \|v\| = 1\}.$$

Recordemos que dos normas en un espacio vectorial de dimensión finita (en este caso, $End(V)$), son siempre equivalentes, así que en realidad podemos tomar cualquier otra norma (como por ejemplo

$$\|T\|_{max} = \max\{|T_{ij}| : i, j = 1, \ldots n\}$$

donde T_{ij} son las entradas de la matriz de T, en alguna base E de V).

Lema 2.23. *Sea $T : V \to V$, tal que cada autovalor t_i de T tiene multiplicidad m_i. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que si $\|R - T\| < \delta$, entonces para todo i, R tiene exactamente m_i autovalores en $B_\varepsilon(t_i) = \{\lambda : |\lambda - t_i| < \varepsilon\}$.*

Demostración. Los autovalores de T son las raíces del polinomio

$$p(\lambda) = \det(T - \lambda 1),$$

mientras que los de R son las raíces de $q(\lambda) = \det(R - \lambda 1)$. Es claro que si T es cercana a R, sus entradas en cualquier base también, y entonces su polinomio característico también. Luego basta tomar δ suficientemente pequeño para que las raíces de q estén cerca de las de p, y obtener el resultado enunciado. ◻

Corolario 2.24. *Sean $p, n \in \mathbb{N}_0$, entonces el conjunto $H_{p,n}$ de las transformaciones lineales $T : V \to V$ autoadjuntas con $p(T) = p, n(T) = n$ es abierto y cerrado en el conjunto de todas las transformaciones autoadjuntas de rango $r = p + n \leq \dim(V)$. En particular $H_{k,l}$ con $k + l = r$ describe las componentes conexas del conjunto de todas las transformaciones autoadjuntas de rango r.*

Demostración. A partir del lema previo tenemos que si R es autoadjunta, de rango r y es cercana T, entonces $p(R) \geq p(T)$ y $n(R) \geq n(T)$. Pero de $r = p(R) + n(R) \geq p(T) + n(T) = r$ deducimos que $p(R) = p(T)$, $n(R) = n(T)$, luego para cada índice fijo (p, n), este conjunto es abierto. Ahora bien el conjunto de todas las transformaciones autoadjuntas de rango r en V es la unión disjunta de los $H_{k,l}$ con k, l recorriendo los posibles valores en $\mathbb{N}_0 \times \mathbb{N}_0$ y la restricción $k + l = r$. Luego el complemento de $H_{p,n}$ también es abierto. ◻

REVISAR TODAS LAS LETRAS

Teorema 2.25 (Ley de inercia de Sylvester - Transformaciones autoadjuntas)**.** *Sean T, R t.l. autoadjuntas. Entonces $p(T) = p(R)$ y $n(T) = n(R)$ si y sólo si existe M inversible tal que $T = MRM^*$.*

Demostración. Si R, T tienen el mismo número de autovalores positivos y negativos, entonces el autoespacio de los autovalores positivos de R tiene la misma dimensión que el de T, y lo mismo vale para el autoespacio de los autovalores negativos, así como para el núcleo. Para T, sus tres autoespacios son ortogonales entre si, y lo mismo vale para R. Si $\{e_1, \ldots, e_n\}$ es una base de autovectores de R y $\{v_1, \ldots, v_n\}$ es una base de autovectores de T (elegida de manera tal que los primeros p son base del autoespacio positivo y los últimos base del núcleo), sea U la transformación ortogonal $Ue_i = v_i$, y sea K dada por $Ke_i = \lambda_i e_i$ para los e_i que no están en el núcleo de R (donde $\lambda_i = \sqrt{t_i r_i^{-1}}$, notar que el cociente tiene el mismo signo, tanto para el primer subespacio de positividad, como para el segundo de negatividad), $Ke_i = e_i$ para los del núcleo. Claramente K es autoadjunta e inversible (Observación 2.9). Ahora consideramos la t.l. definida como $Me_i = \lambda_i v_i$ para los autovalores no nulos de R, y $Me_i = v_i$ para los del núcleo de R. Claramente M es inversible, pero además también es evidente que $M = UK$, luego $M^* = K^* U^* = KU^*$. Ahora bien, $KRKe_i = KR\lambda_i e_i = K\lambda_i r_i e_i = \lambda_i^2 r_i e_i = t_i e_i$ para los vectores que no están en el núcleo, mientras que $KRKe_i = KRe_i = 0$ para los del núcleo. Luego

$$MRM^* v_i = UKRKU^* v_i = UKRKe_i = Ut_i e_i = t_i v_i = Tv_i$$

si v_i no está en el núcleo de T, mientras que $MRM^* v_i = UKRKU^* v_i = UKRKe_i = UKRe_i = 0 = Tv_i$ para los del núcleo de T. Como MRM^* y T coinciden en una base, debe ser $MRM^* = T$.

Supongamos ahora que $T = MRM^*$, escribimos la descomposición polar $M = UA$ con $|M| = A > 0$ y U ortogonal. Es claro que $U^* TU = ARA$, y como $U^* TU$, T tienen exactamente los mismos autovalores, basta probar que R y ARA tienen el mismo índice (cantidad de autovalores positivos y negativos). Escribimos la descomposición espectral de A,

$$A = \sum_{a \in \sigma(A)} a P_a,$$

y como $a > 0$ para todo $a \in \sigma(A)$, podemos definir la curva

$$\alpha(t) = A^t = \sum_{a\in\sigma(A)} a^t P_a, \qquad t \in [0,1].$$

Es claro que $\alpha(1) = A$, $\alpha(0) = 1$ (la identidad de V) pues $\sum P_a = 1$ al ser A inversible. También es claro que α es una curva continua. Entonces

$$t \mapsto A^t R A^t$$

es una curva continua en las transformaciones autoadjuntas, que une R con ARA. Por lo tanto, debe permanecer dentro de una de las componentes conexas, y por el corolario previo, R y ARA tienen el mismo índice. □

Teorema 2.26 (Ley de inercia de Sylvester - formas cuadráticas)**.** *Si T, T^B son las t.l. inducidas por sendos productos internos en V y la misma forma cuadrática Q, ambas transformaciones tienen el mismo rango y signatura.*

Demostración. Recordemos que $T = M^*MT^B$ (Observación 2.14). Para simplificar la notación notamos que M^*M es positiva e invertible, luego $(M^*M)^{-1}$ también lo es, sea $A > 0$ su raíz cuadrada, $A^2 = (M^*M)^{-1}$. Entonces $T^B = (M^*M)^{-1}T = A^2T$, y para cada $\lambda \in \mathbb{R}$, escribimos

$$T^B - \lambda 1 = A^2T - \lambda 1 = A(ATA - \lambda 1)A^{-1}.$$

Esto nos dice que los autovalores de T^B son los mismos que los de la t.l. autoadjunta ATA; como $A > 0$ en particular $A^* = A$ y la ley de inercia de Sylvester para t.l. nos dice que ATA tiene mismo rango y signatura que T. □

Definición 2.27. El lema anterior nos permite hablar de *positividad, negatividad, rango y signatura* de una forma bilineal β, números dados por cualquier representación de β en términos de una transformación autoadjunta (determinada por la elección de una base o un producto interno en V; la transformación va a diferir según la base, pero estos números no).

De la misma manera, el *índice* de beta es el par $(p,n) \in \times\mathbb{N}_0 \times \mathbb{N}_0$ con p la dimensión maximal de positividad de β y n la de negatividad.

3. Funciones cuadráticas y cuádricas

Definición 3.1 (Funciones cuadráticas y cuádricas afines)**.** Sea $Q : V \to \mathbb{K}$ forma cuadrática, denotamos $C_Q = \{v \in V : Q(v) = 0\}$ al conjunto de los ceros de Q. Sean $\varphi \in V^*$ forma lineal y $k \in \mathbb{K}$. Una *función cuadrática* en V está dada por

$$F(x) = Q(x) + 2\varphi(x) + k.$$

Requerimos $Q \neq 0$, y al conjunto de los ceros de F lo denominamos *cuádrica afín*,

$$\mathscr{C} = \mathscr{C}(F) = \{x \in V : F(x) = 0\}.$$

Notar que cuando $\varphi = k = 0$, $\mathscr{C} = C_Q$. Recordando que denotamos $\beta_a = \beta(a,\cdot)$ cuando β es la forma bilineal simétrica asociada a Q, definimos también $\varphi_a = \varphi + \beta_a \in V^*$.

Observación 3.2. Si $d = \dim(V) = 1$, entonces la cuádrica son los ceros de un polinomio cuadrático, luego como $Q \neq 0$, es $F \neq 0$, debe ser $\mathscr{C}$ vacía, un punto doble o dos puntos. Geométricamente, estas cuádricas no tienen más invariantes, y por ello supondremos de aquí en más que $d = \dim(V) \geq 2$.

Observación 3.3 (Múltiplos constantes)**.** En general, para $\lambda \neq 0$, el conjunto de ceros de F y de $F' = \lambda F$ es exactamente el mismo, y es usual decir que hay una única F tal que $\mathscr{C} = \mathscr{C}(F)$ cuando esa sea la única posibilidad.

Definición 3.4 (Cuádricas en espacios con producto interno)**.** Si V es un espacio con producto interno $\langle\cdot,\cdot\rangle$ sobre $\mathbb{R}$ ó $\mathbb{C}$, podemos usar los teoremas de Riesz de representación de formas lineales y bilineales para hallar únicos $T : V \to V$ t.l. autoadjunta y $b \in V$ tales que

$$F(x) = \langle Tx, x\rangle + 2\langle b, x\rangle + k. \tag{7}$$

Notemos que por la definición de cuádrica, debe ser $T \neq 0$.

Si elegimos una base ortonormal $E = \{e_1, \dots e_n\}$ de este espacio, y escribimos $x = \sum_i x_i e_i$, tendremos la escritura en coordenadas

$$F(x) = \sum_{i,j} T_{ij} x_i \overline{x_j} + 2\sum_j b_j \overline{x_j} + k$$

donde el conjugado se puede omitir si el cuerpo es $\mathbb{R}$. Con un pequeño abuso de notación (fijada la base, es decir fijado el sistema de coordenadas), denotaremos a esta expresión $F(x_1, \dots, x_n)$.

Definición 3.5 (Expresión afín de una función cuadrática)**.** Dado $a \in V$, la función F admite una reescritura

$$F(x) = Q(x-a) + 2\varphi_a(x-a) + F(a)$$

que puede pensarse como la expresión afín en el espacio vectorial V_a con origen a. Para probarlo, hay que notar que si uno desconoce φ_a, puede proponer la igualdad y luego operar

$$\begin{aligned} F(x) = Q(x) + 2\varphi(x) + k &= Q(x-a) + 2\varphi_a(x-a) + F(a) \\ &= Q(x) + Q(a) - 2\beta_a(x) + 2\varphi_a(x) - 2\varphi_a(a) + F(a) \end{aligned}$$

para deducir que, como esto debe ser igual a $F(x)$ para todo $x \in V$, primero podemos cancelar los términos constantes tomando $x = 0$. Luego reemplazando x por λx, debe ser

$$\lambda^2 Q(x) + 2\lambda\varphi(x) = \lambda^2 Q(x) - 2\lambda\beta_a(x) + 2\lambda\varphi_a(x),$$

esto sólo es posible si $\varphi = -\beta_a + \varphi_a$.

Observación 3.6. Si $g(x) = Mx + z$ es afín e inversible, y $\widetilde{\mathscr{C}} = g^{-1}(\mathscr{C})$, entonces esta cuádrica está dada por los ceros de $\widetilde{F} = F \circ g$. Un cómputo directo muestra que entonces, si

$$\widetilde{F}(x) = F(g(x)) = \widetilde{Q}(x) + 2\widetilde{\varphi}(x) + \widetilde{k},$$

debe ser

$$\widetilde{Q} = Q \circ M, \qquad \widetilde{\varphi} = \varphi_z \circ M = \beta_z \circ M + \varphi \circ M$$

y finalmente

$$\widetilde{k} = Q(z) + 2\varphi(z) + k = F(z).$$

Es inmediato que entonces $\widetilde{Q}$ está representada por la t.l. autoadjunta $\widetilde{T} = M^*TM$, y que $\widetilde{b} = M^*Tz + M^*b$.

3.1. Cuádricas con centro, puntos singulares y regulares

Definición 3.7. Decimos que $c \in V$ es un *centro* de F si $\varphi_c = 0$ (es la funcional nula en V^*). Decimos que $c \in V$ es un *punto singular* de F si es un centro y además $F(c) = 0$. Definimos la *simetría alrededor de c* $\sigma_c : V \to V$ como la transformación afín $\sigma_c(x) = -x + 2c$.

Notar que σ_c es la simetría alrededor de c, dada por enviar x a $c + (c - x) = 2c - x$. Además σ_c es un isomorfismo con inversa σ_c, porque $\sigma_c \circ \sigma_c = 1$. Veremos en los próximos dos teoremas, que $\mathscr{C}$ es invariante por esta simetría si y sólo si es centro *de cualquier* F que la represente.

Teorema 3.8. *Sea $\mathscr{C}(F)$ una cuádrica, entonces son equivalentes:*

1. *$\mathscr{C} \subset M$ donde $M \subset V$ es una variedad afín propia.*

2. *Todo punto de $\mathscr{C}$ es singular para F.*

3. *Para todo $p \in \mathscr{C}$, $\mathscr{C} = \ker\beta + p$ (en particular $\mathscr{C}$ es una variedad afín propia).*

Demostración. Los casos $\dim(V) = 1$ o $\mathscr{C} = \emptyset$ son triviales. Supongamos que hay un punto v tal que $F(v) = 0$ pero no es singular para F. Entonces $H = \ker(\varphi_v)$ es un subespacio propio. Si la cuádrica está en una variedad afín propia, entonces v está en esa variedad y hay un subespacio $S \subset V$ tal que $\mathscr{C} - v \subset S$. Reescribimos F en su forma afín en v

$$F(x) = Q(x-v) + 2\varphi_v(x-v).$$

Tomamos $z \notin H \cup S \cup C_Q$ (Lema 1.8), entonces $Q(z) \neq 0$ y $\varphi_v(z) \neq 0$. Sea $\lambda = \frac{-2\varphi_v(z)}{Q(z)}$, se tiene

$$\begin{aligned} F(v+\lambda z) &= Q(\lambda z) + 2\varphi_v(\lambda z) = \lambda^2 Q(z) + 2\lambda\varphi_v(z) \\ &= \frac{4\varphi_v^2(z)Q(z)}{Q^2(z)} - \frac{4\varphi_v(z)\varphi_v(z)}{Q(z)} = 0. \end{aligned}$$

Esto dice que $v + \lambda z \in \mathscr{C} \subset v + S$, luego $\lambda z \in S$, y como $\lambda \neq 0$, debe ser $z \in S$, absurdo. Entonces todo punto de la cuádrica es singular para F.

Ahora supongamos que todo punto es singular para F, que $p \in \mathscr{C}$. Si $z \in \ker\beta$, entonces como $F(p) = 0$ y p es centro de F, se tiene

$$F(x) = Q(x-p) + 2\varphi_p(x-p) = Q(x-p).$$

Entonces $F(z+p) = Q(z) = \beta(z,z) = 0$, luego $z + p \in \mathscr{C}$ así que $\ker\beta + p \subset \mathscr{C}$. Tomemos ahora $v \in \mathscr{C}$, entonces

$$\beta_{v-p} = \beta_v + \varphi - (\varphi + \beta_p) = \varphi_v - \varphi_p = 0 - 0 = 0$$

pues todo punto de $\mathscr{C}$ es singular para F. Esto prueba que $v - p \in \ker\beta$ o equivalentemente que $v \in \ker\beta + p$, así que $\mathscr{C} \subset \ker\beta + p$.

Por último, que $\mathscr{C}$ es propia se sigue de que pedimos en la definición de cuádrica que $Q \neq 0$, o equivalentemente que $\ker\beta \neq V$. □

Corolario 3.9 (Caso real)**.** *Sea $\mathscr{C}(F)$ una cuádrica en un $\mathbb{R}$ espacio vectorial V de dimensión n, con producto interno $\langle\cdot,\cdot\rangle$. Son equivalentes:*

1. *$\mathscr{C}(F)$ está dentro de una variedad afín propia.*

2. *Existe una base ortonormal tal que si $p = (p_1, \ldots, p_n) \in \mathscr{C}$ (en esas coordenadas), entonces la expresión de F es*
$$F(x_1, \cdots, x_n) = \sum_{i=1}^{n} \alpha_i (x_i - p_i)^2$$
donde todos los α_i no nulos tienen el mismo signo.

Demostración. Si la expresión de F es la mencionada, con $\alpha_i > 0$ para $i = 1, \ldots r$, entonces el conjunto de ceros de F es exactamente $(p_1, p_2, \ldots, p_r, x_{r+1}, \ldots, x_n)$ con los x_j $(j > r)$ libres, que es una variedad afín propia. Lo mismo ocurre si los α_i no nulos son negativos.

Recíprocamente, si el conjunto de ceros de F es una variedad afín propia, entonces por el teorema anterior todo punto de $\mathscr{C}$ es singular y podemos escribir $F(x) = Q(x-p) = \langle T(x-p), x-p\rangle$ por el Teorema de Riesz, con T autoadjunta. Si $\{v_i\}_{i=1,\ldots n}$ es una base ortonormal de autovectores de T, es claro que $T(x-p) = T\sum_i (x_i - p_i)v_i = \sum_i \alpha_i(x_i - p_i)v_i$, y de allí es inmediata la expresión para F enunciada en el item 2. de este lema; notemos además que derivando llegamos a

$$\nabla F(x_1, \ldots, x_n) = (2\alpha_1(x_1 - p_1), \ldots, 2\alpha_r(x_r - p_r), 0, \ldots, 0)$$

donde suponemos que el núcleo de T está generado por $\{v_{r+1},\cdots,v_n)$. Si existieran $\alpha_i\alpha_j<0$, digamos $\alpha_i>0,\alpha_j<0$, tomamos

$$q=(\sqrt{-\alpha_j}+p_i)v_i+(\sqrt{\alpha_i}+p_j)v_j=x_iv_i+x_jv_j$$

y notamos que

$$F(q)=\alpha_i(\sqrt{-\alpha_j})^2+\alpha_j(\sqrt{\alpha_i})^2=-\alpha_i\alpha_j+\alpha_j\alpha_i=0$$

luego $q\in\mathscr{C}$. Pero $\nabla F(q)=(\ldots,2\alpha_i\sqrt{-\alpha_j},\ldots,2\alpha_j\sqrt{\alpha_i},\ldots)$ debe ser nulo porque todo punto de $\mathscr{C}$ es singular por el teorema anterior, contradicción. $\square$

En las condiciones del teorema anterior, trasladando la cuádrica al origen $0\in V$, debe ser $F(x)=Q(x)$ con la condición adicional que su conjunto de ceros es un subespacio propio. Por ejemplo

$$F(x_1,\ldots,x_n)=x_1^2+x_2^2=0$$

en $\mathbb{R}^n$, con $n\geq 2$.

Teorema 3.10. *Dada una cuádrica $\mathscr{C}=\mathscr{C}(F)$. Entonces $c\in V$ es centro de F si y sólo si $\sigma_c(\mathscr{C})=\mathscr{C}$.*

Demostración. Supongamos primero que $\varphi_c=0$, entonces de la reescritura afín de F en V_c, tenemos para cada $x\in\mathscr{C}$

$$\begin{aligned}F(\sigma_c(x))&=Q(\sigma_c(x)-c)+k=Q(2c-x-c)+k=Q(c-x)+k\\&=Q(x-c)+k=F(x)=0\end{aligned}$$

luego $\sigma_c(\mathscr{C})\subset\mathscr{C}$ pero aplicando nuevamente σ_c y usando que es su propia inversa, se tiene que σ_c deja al conjunto cuádrica invariante.

Supongamos ahora que σ_c deja el conjunto de ceros de F invariante. Para todo $x\in V$ tal que $Q(x-c)+2\varphi_c(x-c)+k=0$, se tiene

$$0=F(\sigma_c(x))=Q(c-x)+2\varphi_c(c-x)+k.$$

Como $Q(x-c)=Q(c-x)$, esto prueba que $\varphi_c(c-x)=0$ en los ceros de F, es decir $\varphi_c|_{\mathscr{C}-c}=0$. Tenemos que ver que φ_c es idénticamente cero en todo V; trasladando al origen (trabajando en V_c) podemos suponer que $c=0$ y que $\varphi|_{\mathscr{C}}=0$. Supongamos que $\varphi\neq 0$, sea $H=ker(\varphi)$; dado $v\in\mathscr{C}$ tenemos $\varphi(v)=0$ luego $\mathscr{C}\subset H$ así que por el Teorema 3.8, $\mathscr{C}$ es una variedad afín propia y todo punto de $\mathscr{C}$ es singular para F. Dado $x\in\mathscr{C}$ tenemos $-x=\sigma_c(x)\in\mathscr{C}$ por hipótesis, y entonces $0=\frac{1}{2}x+\frac{1}{2}(-x)\in\mathscr{C}$. Entonces, como 0 es punto singular de $\mathscr{C}(F)$, debe ser $\varphi=0$ idénticamente, contradicción. $\square$

El teorema previo nos dice que la definición de centro es intrínseca al conjunto cuádrica y no a la F que lo representa (ya que puede haber distintas F que representen la misma cuádrica).

Definición 3.11. El conjunto de centros de una cuádrica $\mathscr{C}$ lo denotamos $centr(\mathscr{C})$, y también $\mathscr{C}_c$. Recordemos que los *puntos singulares* de $\mathscr{C}$ son los centros de $\mathscr{C}$ que estén sobre la cuádrica, esto es $\mathscr{C}_s=\mathscr{C}\cap\mathscr{C}_c$.

Observación 3.12. Notemos que si $x_t=p+tv$ es la recta que pasa por p con dirección v, derivando $F(x_t)$ respecto de t (y evaluando en $t=0$) obtenemos

$$DF_p(v)=DQ_p(v)+2\varphi(v)=2\beta_p(v)+2\varphi(v)=2\varphi_p$$

por lo discutido en la Observación 1.4, es decir $DF_p=2\varphi_p$. Entonces c es un centro de la cuádrica si y sólo si la diferencial de F se anula en c, y es un punto singular si esto ocurre sobre la cuádrica (y esto último ocurre si y sólo si $F(c)=0$).

Notemos también que si V es un espacio con producto interno, entonces

$$\varphi_p = \beta_p + \varphi = \langle Tp, \cdot \rangle + \langle b, \cdot \rangle = \langle Tp + b, \cdot \rangle$$

donde $T : V \to V$, $b \in V$ son como en (7). En ese caso, la diferencial de F de identifica con el gradiente de F, esto es

$$DF_p \simeq \nabla F(p) = 2(Tp + b)$$

o más precisamente $DF_p = \langle \nabla F(p), \cdot \rangle$. Luego un punto p es centro de $\mathscr{C}$ si y sólo si el gradiente de F se anula en p, y es singular si y sólo si esto ocurre sobre la cuádrica.

Definición 3.13 (Puntos regulares, plano tangente)**.** Los puntos no singulares de $\mathscr{C}$ se denominan *puntos regulares*. Por el teorema de la función implícita, como F es C^∞, en un entorno de un punto regular la cuádrica es el gráfico de una función ψ. Luego podemos definir el *plano tangente* a $\mathscr{C}$ en p como el hiperplano afín que es ortogonal a $\nabla F(p) = 2(Tp + b)$ y que pasa por p. Equivalentemente, $T_p\mathscr{C}$ es el núcleo de φ_p, desplazado hasta p), con ecuación

$$\langle Tp + b, x \rangle - \langle Tp + b, p \rangle = 0.$$

Problema 3.14. Sea $\mathscr{C}$ la cuádrica dada por los ceros de la función $F(x) = \langle Tx, x \rangle + 2\langle b, x \rangle + k$.

1. Si $c \in V$ es centro de $\mathscr{C}$, entonces $F(x) = Q(x - c) + F(c)$. Probar que la cuádrica tiene *algún* punto singular si y sólo si $F(c) = 0$.

2. Considerar el sistema afín $Tx + b = 0, \langle b, x \rangle + k = 0$. Probar que $p \in V$ es solución del sistema si y sólo si es un punto singular de la cuádrica. Concluir que el conjunto de puntos singulares de $\mathscr{C}$ es una variedad afín.

Gracias a que tener centro es un propiedad del conjunto $\mathscr{C}$ y no de alguna F que lo define, tenemos las definiciones siguientes:

Definición 3.15 (Clasificación de cuádricas)**.** Diremos que una cuádrica $\mathscr{C}$ es *singular* si tiene algún punto singular, y que es *regular* si no tiene ninguno. Una cuádrica es *con centro* si tiene algún centro.

Diremos que dos cuádricas $\mathscr{C}, \widetilde{\mathscr{C}}$ son *afinmente equivalentes* (o que están en la misma clase afín) si existe alguna transformación afín inversible $g(x) = Mx + a$ tal que $g(\mathscr{C}) = \widetilde{\mathscr{C}}$.

Observación 3.16. Notamos que si $g = Mx + a$ es afin e inversible ($M : V \to V$ t.l. inversible), entonces

$$\begin{aligned}\sigma_{g(c)}(g(x)) &= 2(Mc + a) - (Mx + a) = 2Mc + 2a - Mx - a \\ &= M(2c - x) + a = g(\sigma_c(x)),\end{aligned}$$

luego $g \circ \sigma_c = \sigma_{g(c)} \circ g$ y con esto se puede ver que

$$g(\sigma_c(\mathscr{C})) = \sigma_{g(c)}(g(\mathscr{C}))$$

y que c es un centro de $\mathscr{C}$ si y sólo si $g(c)$ es un centro de $g(\mathscr{C})$, lo que prueba que *ser cuádrica con centro es invariante afín*, y también ser singular es un invariante afín (lo mismo ocurre con ser regular). Notar que F define $\mathscr{C}$ si y sólo si $F \circ g^{-1}$ define $g(\mathscr{C})$.

§ En esta sección entonces hemos probado entonces que toda cuádrica con centro $c \in V$ puede ser descripta, luego de una traslación, como el conjunto de ceros de

$$Q(x) + k = 0$$

donde $k = 0$ si y sólo si la cuádrica con centro es singular. Por la Ley de inercia de Sylvester (Teorema 2.26), el índice $(p, n) \in \mathbb{N}_0 \times \mathbb{N}_0$ de la forma cuadrática Q (Definición 2.27) está unívocamente determinado por la clase afín de la cuádrica $\mathscr{C}$.

3.2. Cuádricas sin centro

Comenzamos por definir la noción de *vértice* que será esencial para llevar las cuádricas sin centro a una forma canónica.

Para esta sección, es conveniente fijar un producto interno, entonces la forma cuadrática Q estará representada por medio de una t.l. autoadjunta T, y la forma lineal por un vector $b \in V$,

$$F(x) = \langle Tx, x\rangle + 2\langle b, x\rangle + k.$$

Respecto de este producto interno, podemos identificar $1/2 DF_v = \varphi_v$ con el vector gradiente $\nabla F(v)$. Luego la condición de regularidad se traduce en pedir que $\nabla F(v) \neq 0$.

Definición 3.17. Sea $\mathscr{C} = \mathscr{C}(F) \subset V$ una cuádrica, diremos que $v \in \mathscr{C}$ es un *vértice* de $\mathscr{C}$ si $0 \neq \nabla F(v) = 2(Tv + b) \in \ker \beta = \ker T$.

Ejemplo 3.18. Veamos en un ejemplo simple qué representan los vértices: sea $F(x,y,z) = y - z^2$ y consideremos $\mathscr{C} = \mathscr{C}(F)$. Representada gráficamente, se trata de una parábola en el plano yz, desplazada a lo largo del eje de las x para formar una "canaleta".

En este caso es claro que, como $Q(x,y,z) = -z^2$, se tiene que $\ker \beta = \ker T$ es ortogonal a la recta generada por $(0,0,1)$.

Por otro lado $\nabla F(x,y,z) = (0,1,-2z)$ que pertenece a $\ker \beta$ si y sólo si $z = 0$. Entonces como $y = z^2$, también debe ser $y = 0$, y el conjunto de vértices de esta cuádrica está dado por todos los puntos

$$\mathscr{C}_c = \{(x,0,0) : x \in \mathbb{K}\},$$

que es donde la superficie se apoya en el plano $z = 0$.

Observación 3.19. Como $V = \ker T \oplus \operatorname{Ran}(T)$ y ambos subespacios son ortogonales, el vector b que induce la forma lineal φ tiene una descomposición única (con $a \in \ker T$)

$$b = a + Tz_0. \tag{8}$$

Afirmamos que $\mathscr{C}$ tiene centro si y sólo si $a = 0$ (o sea si y sólo si $b \in \operatorname{Ran}(T)$). En efecto, si $a = 0$ entonces $b = Tz_0 \in \operatorname{Ran}(T)$ y

$$\varphi_{-z_0} = \varphi + \beta_{-z_0} = \langle b, \cdot\rangle + \langle T(-z_0), \cdot\rangle = \langle b - Tz_0, \cdot\rangle = 0$$

así que $-z_0 \in V$ es centro de $\mathscr{C}$. Recíprocamente, si $c \in V$ es centro de $\mathscr{C}$, entonces $\varphi + \beta_c = 0$ o equivalentemente $b + Tc = 0$, pero $a + Tz_0 + Tc = 0$ sólo es posible si $a = 0$ y $T(z_0 + z) = 0$, ya que rango y núcleo de T están en suma directa.

Volviendo a los vértices, en el próximo lema usamos esta descomposición del vector b para dar condiciones equivalentes, que identifican un $v \in \mathscr{C}$ tal que $\nabla F(v) \in \ker \beta$, pero sin tener en cuenta la condición $\nabla F(v) \neq 0$.

Lema 3.20. *Sea $v \in V$ y $\mathscr{C}(F)$ una cuádrica. Son equivalentes:*

1. $\nabla F(v) \in \ker T$ *y* $F(v) = 0$.
2. $T(Tv + b) = 0$ *y* $F(v) = 0$.
3. $\beta_{\nabla F(v)} = 0$ *y* $F(v) = 0$.
4. $v \in V$ *es solución del sistema afín*

$$(d) = \begin{cases} Tx = a - b & (d1) \\ \langle a + b, x\rangle = -k & (d2) \end{cases}$$

Demostración. Las tres primeras condiciones son equivalentes de manera evidente. Veamos que son equivalentes con la cuarta. Supongamos primero que $Tv+b = Tv+a+Tz_0 = T(v+z_0)+a \in \ker T$ (por la descomposición (8)), entonces como $a \in \ker T$ debe ser $T(v+z_0) \in \ker T$, pero como el rango y el núcleo de T están en suma directa, debe ser $0 = T(v+z_0) = Tv+Tz_0 = Tv+b-a$. Luego $Tv = a-b$, es decir v verifica $(d1)$. Además

$$\begin{aligned}\langle a+b, v\rangle + k &= \langle (a-b), v\rangle + 2\langle b, v\rangle + k \\ &= \langle Tv, v\rangle + 2\langle b, v\rangle + k = F(v) = 0\end{aligned}$$

pues $v \in \mathscr{C}$; luego v verifica también $(d2)$.
Recíprocamente, supongamos que v verifica las ecuaciones (d). Como verifica $(d1)$, tenemos

$$F(v) = \langle (a-b), v\rangle + 2\langle b, v\rangle + k = \langle a+b, x\rangle + k = 0$$

porque también verifica $(d2)$, luego $v \in \mathscr{C}$. Por otro lado de $(d1)$ tenemos $T^2v = Ta - Tb = -Tb$ puesto que $a \in \ker T$, luego $T(Tv+b) = 0$. □

Definición 3.21. Diremos que una cuádrica *tiene vértice* si existe algún vértice para $\mathscr{C}$.

Observación 3.22. Si $v \in \mathscr{C}$ es vértice, y $\widetilde{v} = g^{-1}(v) = M^{-1}(v-z)$, entonces $\widetilde{v} \in \widetilde{\mathscr{C}}$, pero a partir de las expresiones de la Observación 3.6, se tiene

$$\widetilde{T}(\widetilde{T}\widetilde{v} + \widetilde{b}) = M^*TMM^*(Tv+b).$$

Entonces no necesariamente $g^{-1}(v)$ es vértice de $g^{-1}\mathscr{C}$. Sin embargo, esto resulta cierto si M es ortogonal como puede verse en la ecuación de arriba, ya que en ese caso $M^*M = MM^* = 1$.

Afirmamos que tener vértice, sin embargo, *si* es un invariante afín. Esto es consecuencia de la Observación 3.16 y del siguiente teorema.

Teorema 3.23. *Si $\mathscr{C}$ es una cuádrica entonces $\mathscr{C}$ tiene centro si y sólo si no tiene vértice.*

Demostración. Por la Observación 3.19, basta probar que $\mathscr{C}$ tiene vértice si y sólo si $a \neq 0$. Si $\mathscr{C}$ tiene un vértice $v \in \mathscr{C}$, entonces $0 \neq \nabla F(v) = 2(Tv+b) = 2a$ por $(d1)$, luego $a \neq 0$. Supongamos ahora que $a \neq 0$, entonces $a+b$ tampoco puede ser cero ya que

$$\|a+b\|^2 = \|2a+Tz_0\|^2 = 4\|a\|^2 + \|Tz_0\|^2 \geq 4\|a\|^2 > 0$$

pues $a \neq 0$ (núcleo y rango de T son ortogonales). Entonces observamos que $(d2)$ tiene solución, que es de hecho un hiperplano afín. Por otro lado como $a-b = -Tz_0$, también $(d1)$ tiene, separadamente, solución, que es un subespacio afín. Si no hubiera solución de (d), estos subespacios serían paralelos, o equivalentemente, tendriamos que las soluciones del sistema homogéneo $Tx = 0$ serían un subespacio de las soluciones del sistema homogéneo $\langle a+b, x\rangle = 0$. Pero como $Ta = 0$, debería ser $\langle a+b, a\rangle = 0$, esto es $2\|a\|^2 = \langle 2a+Tz_0, a\rangle = 0$, y esto no es posible. Así que el sistema afín (d) tiene solución; como $Tv+b = a \neq 0$ entonces v es un vértice de $\mathscr{C}$. □

Respecto de las posibles funciones cuadráticas que determinana la misma cuádrica, tenemos un resultado interesante que requiere un lema previo (comparar con el Lema 1.7):

Lema 3.24 (Funciones cuadráticas reducibles)**.** *Si $\mathscr{C} = \mathscr{C}(F)$ contiene un hiperplano afín $H = \ker(f)$, entonces existe g forma afín tal que $F = fg$.*

Demostración. Tomando $p \in H$ y trabajando en V_p podemos suponer que H es un hiperplano y que f es una forma lineal. Sea $v \notin H$, escribimos $x \in V$ como $x = \tilde{w} + tv$, con $\tilde{w} \in H$. Vemos que $f(x) = tf(v)$, esto es $t = f(x)/f(v)$. Ahora bien como $F(0) = 0$, tenemos $F = Q + 2\varphi$ con $Q(z) = \beta(z,z)$ no idénticamente nula, luego

$$\begin{aligned}
F(x) = F(\tilde{w} + tv) &= Q(\tilde{w} + tv) + 2\varphi(\tilde{w} + tv) \\
&= Q(\tilde{w}) + t^2 Q(v) + 2t\beta(\tilde{w}, v) + 2\varphi(\tilde{w}) + 2t\varphi(v) \\
&= t^2 Q(v) + 2t\beta(x - tv, v) + 2t\varphi(v) + Q(\tilde{w}) + 2\varphi(\tilde{w}) \\
&= t^2 Q(v) + 2t\beta(x - tv, v) + 2t\varphi(v) \\
&= t^2 Q(v) + 2t\beta(x, v) - 2t^2 Q(v) + 2t\varphi(v) \\
&= t\,(2\beta(x, v) + 2\varphi(v) - tQ(v)) \\
&= \frac{f(x)}{f(v)} \left[2\beta(x, v) + 2\varphi(v) - \frac{f(x)}{f(v)} Q(v) \right]
\end{aligned}$$

con lo cual $F(x) = f(x)g(x)$ con g forma afín. □

Decimos que una tal función cuadrática es *reducible*; notemos que la hipótesis general $Q \neq 0$ (que garantiza también $\beta \neq 0$) es relevante porque en caso contrario nos quedaría $F = \lambda f$ lineal.

Ejemplo 3.25. Consderemos la cuádrica $\mathscr{C} \subset \mathbb{R}^3$ definida como $\mathscr{C} = \{x^2 + y^2 = 0\} = \{2x^2 + 3y^2 = 0\}$. Notemos que $\mathscr{C}$ está definida por dos funciones cuadráticas distintas $F(x,y,z) = x^2 + y^2$ y también $F'(x,y,z) = 2x^2 + 3y^2$, y de hecho hay infinitas F (que no son una múltiplo constante de la otra) que dan la misma cuádrica. En este caso, la cuádrica es una recta, que es una variedad de dimensión $1 = 3 - 2$, donde $d = 3$ es la dimensión del espacio vectorial ambiente $V = \mathbb{R}^3$.

El siguiente teorema nos dice que esto sólo puede pasar por la diferencia de dimensiones:

Teorema 3.26 (Teorema de unicidad)**.** *Sea $\mathscr{C} \subset V$ con $dim(V) = d$ una cuádrica. Supongamos que $\mathscr{C}$ es el conjunto de ceros de las funciones cuadráticas F, F', y $\mathscr{C}$ no está contenido en una variedad lineal de dimensión $d - 2$. Entonces existe $\lambda \neq 0$ tal que $F' = \lambda F$.*

Demostración. Vamos a dar las ideas importantes de la prueba, dejando algunos puntos como ejercicio para el lector. Probemos primero el teorema en dimensión $d = 2$. Supongamos que $\mathscr{C}$ contiene una recta $L = \ker(f)$ con f forma afín, por el lema previo vemos que $F(x) = f(x)g(x)$ con g forma afín. Ahora bien, la misma cuenta nos dice que $F' = fg'$ y como F, F' tienen los mismo ceros debe ser $g' = \lambda g$ para algún lambda, con lo cual $F' = \lambda F$. Ahora supongamos que $\mathscr{C}$ está contenida en una recta, entonces por el Teorema 3.8 es una variedad lineal, como no puede ser un punto $\mathscr{C}$ por la hipótesis, debe ser una recta, luego contiene una recta y por el caso anterior $F' = \lambda F$ nuevamente. Podemos suponer entonces que la cuádrica no contiene ni está contenida en ninguna recta. Existen entonces cinco puntos $A, B, C, D, E \in \mathscr{C}$ de lo cuales no hay tres alineados (Ejercicio). Sea L la recta por A,B, tenemos que $L \cap \mathscr{C} = \{A, B\}$ (Ejercicio: una recta corta una cuádrica en $0, 1, 2$ puntos o en toda la recta). Si restringimos F, F' a la recta L, tenemos dos funciones cuadráticas en dimensión 1; esto es dos polinomios cuadráticos. Como los ceros de ambas son A, B con $A \neq B$, debe ser $F'|_L = \lambda F|_L$ para algún $\lambda \neq 0$ (Ejercicio). Sea

$$G = F' - \lambda F = Q' - \lambda Q + 2\varphi' - 2\lambda\varphi + k' - \lambda k$$

definida en V, sea $\mathscr{K}$ su conjunto de ceros. Es claro que $L \subset \mathscr{K}$ y si fuese $Q' - \lambda Q \neq 0$ tendríamos que G es reducible, luego $\mathscr{K} = L \cup L_1$ pero los puntos C, D, E no están en L, asíq eu deben estar en L_1. Pero en ese caso estarían alineados, contradicción: debe ser $Q' - \lambda Q = 0$ y G es lineal afín. Entonces $\mathscr{K}$ es una recta o es todo V; pero no puede ser una recta porque $\mathscr{C} \subset \mathscr{K}$ y $\mathscr{C}$ no estaba contenida en ninguna recta. Así que $\mathscr{K} = V$ y $G = 0$, lo que prueba que $F' = \lambda F$. Supongamos ahora que $dim(V) = d \geq 3$. Si $\mathscr{C}$ está contenida en una variedad lineal propia, es una variedad lineal propia por el Teorema 3.8, y por la hipótesis (no está contenida en una variedad de dimensión $d - 2$), $\mathscr{C}$ tiene que ser un hiperplano, pero por el lema anterior tenemos y razonando como en dimensión 2 vemos que $F' = \lambda F$ en este caso. Podemos suponer entonces que $\mathscr{C}$ no es una variedad lineal; por el Teorema 3.8 existe un punto no singular $c \in \mathscr{C}$.

Trabajando en V_c, podemos suponer que $c = 0$. Tenemos $F = Q + 2\varphi$, $F' = Q' + 2\varphi'$ con $\varphi \neq 0$, $\varphi' \neq 0$ (porque el origen es un punto regular de la cuádrica). Sea $H = \ker(\varphi)$, no puede ser $\mathscr{C} \subset H$ porque sería toda singular, luego existe $v \in \mathscr{C}$ tal que $v \notin H$, i.e. $\varphi(v) \neq 0$. Tomemos $0 \neq p \in H$ y sea $\pi = gen(p, v)$ que es un espacio vectorial de dimensión 2, y escribimos $x = sp + tv \in S$. Tenemos

$$\begin{aligned} F(x) =& Q(sp+tv) + 2\varphi(sp+tv) = s^2 Q(p) + t^2 Q(v) + 2st\beta(p,v) \\ &+ 2s\varphi(p) + 2t\varphi(v) = H(s,t) \end{aligned}$$

que es un polinomio cuadrático en π, y entonces es claro que $\mathscr{C}_\pi = \mathscr{C} \cap \pi$ es una cuádrica allí (Ejercicio: probar que en general, la intersección de una cuádrica con una variedad afín es una cuadrica en ese espacio vectorial). Tenemos una expresion similar para F', y sabemos que tanto F como F' se anulan en $\mathscr{C}_\pi$, debe ser $F'|_\pi = \lambda_p F|_\pi$ para algún $\lambda_p \neq 0$ por lo probado anteriormente (a priori dependiendo de $p \in H$). En particular

$$\begin{aligned} Q'(p) &= \lambda_p Q(p), & \beta'(p,v) &= \lambda_p \beta(p,v), \\ Q'(v) &= \lambda_p Q(v), & \varphi' &= \lambda_p \varphi. \end{aligned}$$

Ahora bien, tenemos $\varphi'(p) = \lambda_p \varphi(p) = 0$, y entonces $p \in \ker(\varphi')$. Como este argumento se puede hacer para cualquier $p \in H$, vemos que $H \subset \ker(\varphi')$. Como $\varphi' \neq 0$ debe ser $\ker(\varphi') = H = \ker(\varphi)$ con lo cual existe $\lambda \neq 0$ (independiente de p) tal que $\varphi = \lambda\varphi'$. Vemos que $\lambda\varphi(v) = \varphi'(v) = \lambda_p \varphi(v)$; cancelando $\varphi(v)$ vemos que $\lambda_p = \lambda$ (esto para todo $p \in H$) y de allí tenemos las cuatro identidades de arriba pero para λ fijo. Entonces si $z \in V$, escribimos $z = p + tv$ para algún $p \in H$ y algún $t \in \mathbb{K}$, y tenemos

$$\begin{aligned} F'(z) =& Q'(p+tv) + 2\varphi'(p+tv) \\ =& Q'(p) + t^2 Q'(v) + 2t\beta'(p,v) + 2\varphi'(p+tv) \\ =& \lambda Q(p+tv) + 2\lambda\varphi(p+tv) = \lambda F(z), \end{aligned}$$

lo que termina de probar que $F' = \lambda F$ en V. □

3.3. Clasificación ortogonal

En esta sección probamos los teoremas de clasificación en clases ortogonales de las cuádricas.

Teorema 3.27 (Clasificación normal Euclidea de cuádricas sobre $\mathbb{R}$). *Sea $E = \{e_1, \ldots, e_d\}$ una base de V, sea $\mathscr{C} \subset V$ una cuádrica y sea*

$$F(x) = \sum_{i,j} T_{ij} x_i x_j + 2\sum_j b_j x_j + k = 0$$

donde $x = \sum_i x_i e_i \in V$, y para $i, j = 1, \ldots d$, los coeficientes $T_{ij}, b_j, k \in \mathbb{R}$ (al menos un T_{ij} no nulo). Supongamos que $\mathscr{C} = \mathscr{C}(F)$. Existe entonces una transformación ortogonal $U : V \to V$ tal que

1. *Si $c \in V$ es un punto singular de $\mathscr{C}$, la cuádrica es equivalente por el isomorfismo ortogonal afín $g(x) = Ux + c$ a la cuádrica $\mathscr{C}(F_\lambda)$, donde fijada una d-upla $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ se tiene*

$$F_\lambda(x) = (F \circ g)(x) = \sum_{i=1}^{d} \lambda_i x_i^2. \tag{9}$$

   Estas funciones cuadráticas están en clases ortogonales disjuntas. Si $\mathscr{C}$ no está dentro de una variedad afín de dimensión $d-2$, la cuádrica determina unívocamente (salvo un múltiplo constante) al polinomio F_λ. Cuando $\mathscr{C}$ está dentro de una tal variedad afín, no hay unicidad pero podemos suponer que $\lambda_i \geq 0$ para todo i, y $\mathscr{C}$ es una variedad afín.

2. *Si $c \in V$ es centro de $\mathscr{C}$ y la cuádrica es no singular, entonces es equivalente por el isomorfismo ortogonal afín $g(x) = Ux + c$ a una y sólo una (salvo múltiplo constante) de las cuádricas determinadas unívocamente por la d-upla $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$.*

$$\sum_{i=1}^{d} \lambda_i x_i^2 + 1 = 0. \tag{10}$$

3. *Si $\mathscr{C}$ no tiene centro, entonces tiene un vértice $v \in \mathscr{C}$, y la cuádrica es equivalente por el isomorfismo ortogonal afín $g(x) = Ux + v$ a una y sólo una (salvo múltiplo constante) de las cuádricas*

$$\sum_{i=1}^{d-1} a_i x_i^2 + x_d = 0, \tag{11}$$

determinada por la d-1-upla a_i, siempre que $|a_1| \geq |a_2| \geq \cdots \geq |a_{d-1}|$ y que $a_1 > 0$.

Las expresiones (10), (11) se conocen como *forma normal Euclídea* de una cuádrica. Observemos que por el cambio de variables aplicado, el centro de la cuádrica en esta forma canónica es $0 \in V$. También vemos que para el v elegido en las sin centro, la dirección de $\nabla F(v) \in \ker T$ se transforma en el vector $(0, 0, \ldots, 0, 1)$, y puede el lector verificar que el conjunto de vértices es el subespacio $x_{r+1} = x_{r+2} = \cdots = x_d = 0$ donde $(a_i)_{i=1,\ldots,r}$ son los coeficientes no nulos.

Demostración. Hemos probado que se puede separar a las cuádricas en dos clases afines, en particular clases ortogonales: las cuádricas con centro y las sin centro. Supongamos primero que $\mathscr{C}$ tiene centro $c \in V$, entonces está dada por una función cuadrática $F(x) = Q(x - c) + F(c)$. Recordemos que $\mathscr{C}$ es singular si y sólo si $F(c) = 0$ (Problema 3.14); si $\mathscr{C}$ es no singular, dividiendo por $F(c)$ tenemos que $\mathscr{C}$ es el conjunto de ceros de $F(x) = Q(x - c) + 1$ (donde hemos reemplazado Q por un múltiplo). Sea $B = \{v_1, \ldots, v_d\}$ una base de V que diagonaliza T, la t.l. autoadjunta determinada por Q, sean $\{\lambda_i\}_{i=1,\ldots,d}$ los autovalores de T, y sea $U : V \to V$ la transformación ortogonal tal que $Ue_i = v_i$. Tomamos $g(x) = Ux + c$ isomorfismo ortogonal afín. Consideremos $\tilde{F} = F \circ g$, entonces $\mathscr{C}$ es ortogonalmente equivalente al conjunto de los ceros de $\tilde{F}$, es decir $\tilde{\mathscr{C}} = g^{-1}(\mathscr{C})$. Veamos que $\tilde{F}$ tiene la forma deseada. Para ello observamos que si $x = \sum x_i e_i$ entonces $Ux = \sum x_i v_i$ y además $TUx = \sum_i \lambda_i x_i v_i$, luego

$$\tilde{F}(x) = F(g(x)) = Q(Ux) + \mu = \langle TUx, Ux \rangle + \mu = \sum_i \lambda_i x_i^2 + \mu,$$

donde $\mu = 1$ cuando $\mathscr{C}$ es regular y $\mu = 0$ si es singular, como afirma el enunciado del Teorema. Si dos funciones cuadráticas de estas son ortogonalmente equivalentes, debe ser $\widetilde{T} = U^* T U$ (Observación 3.6), luego los autovalores de $\widetilde{T}$ y los de T son exactamente los mismos, y la condición de orden no permite permutaciones no triviales (y además $\mu = 0$ si y sólo si $\mathscr{C}$ es singular, que es un invariante afín). Si $\mathscr{C}$ es regular o es singular pero no está en una variedad afín de dimensión $d - 2$, el teorema de unicidad nos dice que F_λ es única para ese conjunto de ceros (salvo mútiplo constante). Cuando $\mathscr{C}$ está dentro de una variedad afín de dimensión $d - 2$, el Teorema 3.8 y su corolario nos dicen que podemos elegir todos los $\lambda_i \geq 0$ y que la cuádrica es una variedad afín.

Ahora supongamos que $\mathscr{C}$ no tiene centro; por el Teorema 3.23, tiene al menos un vértice $v \in \mathscr{C}$. Como $\nabla F(v) \in \ker T$, el subespacio $S = \{\nabla F(v)\}^\perp$ es invariante para T (verificarlo!). Entonces $T|_S$ se diagonaliza con una base ortonormal de S, y podemos armar una b.o.n. $B = \{v_1, \ldots, v_{d-1}, v_d\}$ de manera tal que $\{v_1, \ldots, v_{d-1}\}$ es b.o.n. de S y $v_d = \nabla F(v) \|\nabla F(v)\|^{-1}$. Ahora denotamos con a_i ($i = 1, \ldots, d-1$) a los primeros $d - 1$ autovalores de T (notar que el último autovalor es nulo pues corresponde a $\nabla F(v)$), y denotamos $a_d = \|\nabla F(v)\|$. Nuevamente, sea U la transformación ortogonal tal que $Ue_i = v_i$, y sea $g(x) = Ux + v$. Como antes, $\mathscr{C}$ es ortogonalmente equivalente al conjunto de los ceros de $\tilde{F}$, donde $\tilde{F} = F \circ g$. Veamos que $\tilde{F}$ tiene la forma deseada, para eso primero notemos que

$$2\varphi_v = 2\varphi + 2\beta_v = 2\langle b + Tv, \cdot \rangle = \langle \nabla F(v), \cdot \rangle = \|\nabla F(v)\| \langle v_d, \cdot \rangle,$$

luego al escribir la forma afín con origen v tenemos

$$F(x) = Q(x-v) + 2\varphi_v(x-v) = Q(x-v) + \|\nabla F(v)\|\langle v_d, x-v\rangle.$$

Dividiendo por $\|\nabla F(v)\|$, resulta que $\mathscr{C}$ es el conjunto de ceros de $F(x) = Q(x-v) + \langle v_d, x-v\rangle$ (nuevamente reemplazando Q por un múltiplo). Como $Tv_d = 0$ y $v_i \perp v_d$ para todo $i < d$

$$\begin{aligned}\tilde{F}(x) = F(g(x)) &= Q(Ux) + \langle v_d, Ux\rangle \\ &= \sum_{i=1}^{d-1} a_i x_i^2 + \langle v_d, \sum_{i=1}^{d} x_i v_i\rangle = \sum_{i=1}^{d-1} a_i x_i^2 + x_d\end{aligned}$$

que es la expresión deseada. Nuevamente, si dos cuádricas son ortogonalmente equivalentes, tenemos la invariancia de los autovalores de T por conjugación con una t.l. ortogonal, lo que determina los a_i unívocamente (con la condición enunciada, pensarlo). Al ser no singular el teorema de unicidad no permite dos expresiones como estas para una misma cuádrica (salvo múltiplos constantes). $\square$

3.4. Clasificación afín

Veamos ahora una lista exhaustiva (y disjunta) de todas las clases de equivalencia afín de las cuádricas, separadas en las que tienen y las que no tienen centro.

Teorema 3.28 (Cuádricas con centro - Clasificación afín sobre $\mathbb{R}$)**.** *Si $\mathscr{C}$ es una cuádrica con centro, existen un único par $(p,n) \in \mathbb{N}_0 \times \mathbb{N}_0$ con $p+n \leq \dim(V) = d$, y una base $B = \{v_1, \ldots, v_d\}$ de V, tal que $\mathscr{C}$ es equivalente (por un isomorfismo afín) a una y sólo una de las cuádricas dada por*

$$\sum_{i=1}^{p} x_i^2 - \sum_{j=p+1}^{p+n} x_j^2 + \mu = 0, \tag{12}$$

donde $x = \sum_i x_i v_i \in V$. Se tiene $\mu = 0$ si y sólo si $\mathscr{C}$ es singular, y $\mu = 1$ si y sólo si $\mathscr{C}$ es regular.

Demostración. Elegimos una base $E = \{e_1, \ldots, e_d\}$ de V y el producto interno que ella induce. La cuádrica $\mathscr{C}$ tiene centro, por lo tanto es ortogonalmente (y entonces, afinmente) equivalente a una de las dadas por (9,10), podemos entonces suponer que está dada por esa ecuación en la base E. Ahora definimos la t.l. inversible $M : V \to V$ de la siguiente forma: para los autovalores positivos de T, $Me_i = \sqrt{\lambda_i}^{-1} e_i$, para los negativos $Me_j = \sqrt{-\lambda_j}^{-1} e_j$ y para los nulos $Me_k = e_k$. Resulta así que $\mathscr{C}$ es afínmente equivalente a una de las dadas en (12). Lo que tenemos que ver, es que dos de ellas están en clases afines disjuntas si $(p_1, n_1) \neq (p_2, n_2)$. Pero esto es inmediato de la Observación 3.6 que relaciona las t.l. autoadjuntas luego de una transformación afín como $\widetilde{T} = M^* T M$, y la Ley de inercia de Sylvester (Teorema 2.25) que dice que estas dos transformaciones tiene el mismo índice (p,n). $\square$

La expresión (12) se conoce como la *forma canónica afín* de la cuádrica $\mathscr{C}$.

En el caso de $\mathbb{C}$ espacios vectoriales, cambiando las segundas n variables x_j (las correspondientes a los coeficientes negativos) por ix_j, tenemos $(ix_j)^2 = -x_j^2$ y entonces se obtiene el corolario:

Teorema 3.29 (Cuádricas con centro - Clasificación afín sobre $\mathbb{C}$)**.** *Si $\mathscr{C}$ es una cuádrica con centro, existen un único $1 \leq r \leq \dim(V) = d$, y una base $E = \{e_1, \ldots, e_d\}$ de V, tal que $\mathscr{C}$ es equivalente (por un isomorfismo afín) a la cuádrica dada por*

$$\sum_{i=1}^{r} x_i^2 + \mu = 0,$$

donde $\mu = 0$ si $\mathscr{C}$ es singular, y $\mu = 1$ si $\mathscr{C}$ es regular.

Veamos ahora como es la clasifación cuando $\mathscr{C}$ no tiene centro; en ese caso sabemos que tiene al menos un vértice, obtenemos así, con una demostración similar a la del teorema previo, invocando la clasificación normal Euclídea de las cuádricas sin centro:

Teorema 3.30 (Cuádricas sin centro - Clasificación afín sobre $\mathbb{R}$)**.** *Si $\mathscr{C}$ es una cuádrica sin centro, existen un único par $(p,n) \in \mathbb{N}_0 \times \mathbb{N}_0$ con $p+n \leq \dim(V)-1 = d-1$, y una base $E = \{e_1, \ldots, e_d\}$ de V, tal que $\mathscr{C}$ es equivalente (por un isomorfismo afín) a la cuádrica dada por*

$$\sum_{i=1}^{p} x_i^2 - \sum_{j=p+1}^{p+n} x_j^2 + x_d = 0,$$

única salvo el intercambio de p,n.

Teorema 3.31 (Cuádricas sin centro - Clasificación afín sobre $\mathbb{C}$)**.** *Si $\mathscr{C}$ es una cuádrica sin centro, existen un único $1 \leq r \leq \dim(V)-1 = d-1$, y una base $E = \{e_1, \ldots, e_d\}$ de V, tal que $\mathscr{C}$ es equivalente (por un isomorfismo afín) a la cuádrica dada por*

$$\sum_{i=1}^{r} x_i^2 + x_d = 0.$$

Los detalles de las demostraciones quedan como ejercicio para el lector.