Capítulo 6. Formas bilineales y cuadráticas

Introducción

El Álgebra Lineal no agota su estudio con las aplicaciones lineales, por importantes que sean; ya tratamos (al estudiar los determinantes) con aplicaciones multilineales. Cuando queremos estudiar Geometría, resulta conveniente usar el producto escalar, que es una aplicación bilineal. Cuando investigamos figuras por sus ecuaciones, las de primer grado describen rectas y planos; otras figuras más complicadas (elipses, parábolas, conos, hiperboloides, ...) requieren ecuaciones de segundo grado, que corresponden a formas cuadráticas. En este capítulo nos vamos a ocupar de ellas.

Definición.- Sea V un espacio vectorial de dimensión finita, n. Una *forma bilineal simétrica* en V es una aplicación $f : V \times V \to \mathbf{R}$ tal que

$f(u+v,w) = f(u,w) + f(v,w)$, $f(\lambda \cdot u, v) = \lambda \cdot f(u,v)$ $\forall u,v,w \in V$, $\forall \lambda \in \mathbf{R}$

$f(u,v) = f(v,u)$, $\forall u,v \in V$

La dos primeras condiciones se pueden resumir así:

$f(\lambda \cdot u + \mu \cdot v, w) = \lambda \cdot f(u,w) + \mu \cdot f(v,w)$

Nota.- A veces se habla de formas bilineales sin exigirles la simetría; en ese caso, la condición $f(u,v) = f(v,u)$, $\forall u,v \in V$ se debe eliminar y poner en su lugar $f(u, v+w) = f(u,v) + f(u,w)$, $f(u, \lambda \cdot v) = \lambda \cdot f(u,v)$ $\forall u,v,w \in V$, $\forall \lambda \in \mathbf{R}$ (o bien $f(w, \lambda \cdot u + \mu \cdot v) = \lambda \cdot f(w,u) + \mu \cdot f(w,v)$). En el caso de las formas simétricas, no es necesario pedirlo expresamente, puesto que se deduce de la simetría y la linealidad en la primera variable.

Las formas bilineales no simétricas tienen un interés muy escaso, y nosotros no las estudiaremos; salvo por un ejemplo que citarmos a continuación, todas las formas bilineales que veamos serán simétricas.

Ejemplos y ejercicios.

$f((x,y),(x',y')) = x \cdot x' + 4x \cdot y' - 3y \cdot x'$ es una forma bilineal en $\mathbf{R}^2$. No es simétrica.

$f((x,y),(x',y')) = 2x \cdot y' + 5x - 3y$ no es bilineal.

$f((x,y),(x',y')) = 3x \cdot x' + 2x \cdot y' + 2y \cdot x' + y \cdot y'$ es una forma bilineal simétrica en $\mathbf{R}^2$.

$f((x,y),(x',y')) = x \cdot x' + y \cdot y'$ es otra forma bilineal simétrica en $\mathbf{R}^2$. El alumno probablemente la conozca ya, puesto que es el producto escalar usual.

Del mismo modo $f((x,y,z),(x',y',z')) = x \cdot x' + y \cdot y' + z \cdot z'$ es una forma bilineal simétrica en $\mathbf{R}^3$ que describe el producto escalar ususal en 3 dimensiones.

$x \cdot x' - x \cdot z' - z \cdot x' - y \cdot z' - z \cdot y' + 2z \cdot z'$ define otra forma bilineal simétrica en $\mathbf{R}^3$.

Matriz de una aplicación bilineal simétrica (respecto de una base)

Las matrices, que ya probaron su utilidad para el estudio y el cálculo con aplicaciones lineales, son también muy convenientes para trabajar con formas bilineales. Veámoslo.

Una forma bilineal simétrica en el espacio vectorial V se conoce tan pronto como se sabe el efecto que tiene sobre los vectores de una base: si conocemos los valores $a_{ij} = f(u_i, u_j)$ cuando $\{u_1, \ldots, u_n\}$ es una base de V, ya conocemos todo sobre la forma bilineal f, puesto que cualquier par de vectores u, v de V se escriben como unas combinaciones lineales

$u = \lambda_1 \cdot u_1 + .. + \lambda_n \cdot u_n$

$v = \mu_1 \cdot u_1 + .. + \mu_n \cdot u_n$

y por la bilinealidad se tendrá

$f(u, v) = \sum \lambda_i \cdot \mu_j \cdot f(u_i, u_j) = \sum \lambda_i \cdot \mu_j \cdot a_{ij}$

Así pues, la matriz $M = (a_{ij})$ contiene toda toda la información sobre la forma bilineal f. Llamaremos a la matriz M "*matriz de f respecto de la base dada*". Precisando:

Definición.- Si $f : V \times V \to \mathbf{R}$ es una aplicación bilineal, B $= \{u_1, .., u_n\}$ es una base de V, entonces *la matriz de f respecto de B* es la matriz $n \times n$ definida por $a_{ij} = f(u_i, u_j)$.

Nótese que para cualquier par de vectores $u, v \in V$, el valor $f(u, v)$ se puede calcular haciendo

$f(u, v) = \sum \lambda_i \cdot \mu_j \cdot f(u_i, u_j) = \sum \lambda_i \cdot \mu_j \cdot a_{ij}$

que no es más que el producto $u_B^T \cdot M \cdot v_B$

Ejemplo.- La matriz de la forma $f((x, y, z), (x', y', z')) = x \cdot x' + x \cdot z' + z \cdot x' - y \cdot z' - z \cdot y'$ en la base canónica de $\mathbf{R}^3$ es

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

Nota importante.- Esta matriz es simétrica, y no por casualidad: la simetría de la forma bilineal se refleja en la matriz. Si considerásemos formas bilineales no simétricas, vendrían representadas por matrices asimétricas.

Efecto de un cambio de base

Si una forma bilineal, f, se expresa en la base B mediante la matriz M, y en otra base C mediante la matriz N, es de esperar que entre las matrices M y N haya una relación estrecha. En efecto, esa relación es fácil de precisar:

Proposición.- Sea P = (B:C) la matriz de paso de C a B. Entonces, la matriz de f respecto de C, N, es el producto $P^T \cdot M \cdot P$ (es decir: un cambio de base en el espacio se traduce en

términos matriciales por una multiplicación a la derecha por la matriz del cambio de base, y una multiplicación a la izquierda por la traspuesta de esa matriz).

Demostración.- La comprobación es inmediata: dados dos vectores cualesquiera, u y v, se tiene $f(u,v) = u_B^T \cdot M \cdot v_B = (P \cdot u_C)^T \cdot M \cdot (P \cdot v_C) = u_C^T \cdot P^T \cdot M \cdot P \cdot v_C$; comparándolo con $f(u,v) = u_C^T \cdot N \cdot v_C$ se concluye.

Definición.- Dos matrices cuadradas del mismo tamaño, M y N, se dicen *congruentes* cuando existe una matriz regular P tal que $N = P^T \cdot M \cdot P$.

Nótese que *matrices congruentes corresponden a una misma forma bilineal expresada en bases diferentes.*

Observación: Esa relación es simétrica, es decir, que los papeles de N y M se pueden intercambiar en la definición. También es transitiva: si M y N son congruentes y también lo son N y J, entonces lo serán igualmente M y J.

Dos matrices congruentes han de tener el mismo rango, puesto que la relación $N = P^T \cdot M \cdot P$ obliga en particular a que M y N sean matrices equivalentes. Podemos, pues, hablar del *rango de una forma bilineal* como el rango de cualquiera de las matrices que la representan.

Por otra parte, la congruencia de matrices es una relación más exigente que la equivalencia, y el rango no es suficiente para caracterizarla. Más adelante estudiaremos la congruencia de matrices simétricas.

Ejemplo.- Las matrices

$$M = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

$$N = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$$

son congruentes, puesto que representan a la misma forma bilineal, en la base canónica y en la base $\{(2,1),(1,-1)\}$ La matriz P es

$$P = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$$

Compruebe que, en efecto, $N = P^T \cdot M \cdot P$.

Conjugación

Definición.- Dada una forma bilineal simétrica, f, en el espacio vectorial V, decimos que el vector u es *conjugado* del vector v respecto de esa forma bilineal si $f(u,v)=0$.

Al ser f simétrica, si u es conjugado de v, v también es conjugado de u, por lo que la relación de conjugación es simétrica a su vez. Decimos simplemente que u y v son vectores conjugados.

Ejemplo.- Sea la forma f cuya matriz en la base canónica de $\mathbf{R}^3$ es

$$A=\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

entonces los vectores $(1,0,1)^T$ y $(2,1,0)^T$ son conjugados. El vector $(1,1,1)$ es conjugado de sí mismo. Tales vectores se llaman *autoconjugados*. Para algunas formas, el único vector autoconjugado es el nulo; otras admiten vectores autoconjugados no nulos.

Proposición.- Si S es un subconjunto de V, el conjunto S^0 formado por los vectores conjugados de todos los de S es un subespacio vectorial de V.

Demostración: Es muy sencilla. Sean u y v dos vectores de S^0, y sean λ y μ dos escalares; para ver que $\lambda\cdot u+\mu\cdot v$ está en S^0 sólo hay que ver que $f(\lambda\cdot u+\mu\cdot v,s)$ es igual a 0 para cualquier $s\in S$, y eso es obvio, pues $f(\lambda\cdot u+\mu\cdot v,s)=\lambda\cdot f(u)+\mu\cdot f(v,s)=\lambda\cdot 0+\mu\cdot 0=0$.

Definición.- El conjunto de los vectores de V que son conjugados de todos los vectores, $\{u\in V : f(u,v)=0,\forall v\in V\}$ se llama *núcleo de la forma bilineal*. De acuerdo con la proposición anterior, es un subespacio vectorial de V, pues no es otra cosa que V^0. Sus ecuaciones en una base B son muy fáciles de escribir; $M\cdot X=0$, siendo M la natriz de f en esa base. De ahí se deduce inmediatamente que la dimensión del núcleo de f es igual a la dimensión de V menos el rango de M (que es el de f), es decir, que la dimensión del núcleo más el rango de la forma bilineal es igual a la dimensión del espacio.

Definición.- Una forma cuadrática, q, se dice *degenerada* cuando su núcleo no es trivial (o sea, cuando hay algún vector que es conjugado de todo el espacio), y se dice *no degenerada* en caso contrario.

Es inmediato que una forma cuadrática no degenerada es aquella que tiene rango máximo, lo que se traduce en que el determinante de su matriz (en alguna base) no se anula.

Ejercicio.- Si S y T son subconjuntos de V y $S\subset T$, compruebe que $T^0\subset S^0$.

Ejercicio.- Si S es un subconjunto de V y $U=L(S)$ es el subespacio vectorial generado por S, compruebe que los conjuntos S^0 y U^0 coinciden.

Formas cuadráticas

Definición.- Si $f : V \times V \to \mathbf{R}$ es una forma bilineal simétrica, la aplicación $q : V \to \mathbf{R}$ definida por $q(u) = f(u, u)$ se denomina *forma cuadrática* asociada a f, y f es la *forma polar* de q.

Observaciones: A veces se define forma cuadrática a partir de una forma bilineal a secas (sin exigirle la simetría). Así, una forma cuadrática viene asociada a muchas formas bilineales. Sin embargo, al exigir simetría ganamos la reciprocidad, puesto que una forma cuadrática sólo está asocia una forma bilineal simétrica (su polar), que se puede recuperar fácilmente a partir de aquella. Es sencillo comprobar que si f es la forma polar de q, entonces $f(u, v) = [q(u + v) - q(u) - q(v)]/2$.

El emparejamiento entre formas cuadráticas y formas bilineales simétricas nos permite traer aquí conceptos de éstas. Así, llamaremos matriz de una forma cuadrática en una base a la de su froma polar, núcleo y rango de la forma cuadrática a los de su polar, y diremos que dos vectores son conjugados respecto de la forma cuadrática cuando lo sean respecto de la polar.

Diagonalización de formas cuadráticas.- Al escribir la expresión de la forma cuadrática en una base como $q(u) = u_B^T \cdot M \cdot u_B$, se observa que es un polinomio homogéneo de segundo grado. Tal fórmula será siempre sencilla, pero lo es especialmente cuando no involucra productos mixtos, sino que se reduce a una suma algebraica de cuadrados (es decir, algo así como $x^2 + 3y^2 - 6z^2$, frente a la más complicada $x^2 + 3xy - y^2 + 5xz - 2z^2$). Puesto que la expresión de una forma cuadrática varía al cambiar de base, cabe pensar que eligiendo ésta adecuadamente se pueda conseguir expresar la forma cuadrática como una suma de cuadrados. Eso es lo que llamamos "diagonalizar una forma cuadrática".

Definición.- Decimos que hemos diagonalizado la forma cuadrática q cuando hemos encontrado una base de V en la cual la matriz de q sea una matriz diagonal. Eso equivale a que la base esté formada por vectores conjugados dos a dos respecto de la forma q. También decimos que hemos expresado q como una suma de cuadrados, porque se tiene $q(u) = d_1 \cdot x_1^2 + \ldots + d_n \cdot x_n^2$, siendo $x_1, \ldots, x_n$ las coordenadas de u en esa base, y $d_1, \ldots, d_n$ los elementos que ocupan la diagonal de la matriz.

En términos de la matriz de q, pasa de ser una matriz M a ser una matriz diagonal $N = P^T \cdot M \cdot P$. Por eso, llamamos *diagonalizar por congruencia la matriz M* a encontrar una matriz regular, P, tal que $P^T \cdot M \cdot P$ sea una matriz diagonal.

Pueden caber dudas acerca de si será posible o no diagonalizar una forma cuadrática. El siguiente teorema las disipa:

Teorema.- Si q es una forma cuadrática en el espacio vectorial V, entonces hay una base de V formada por vectores conjugados respecto de q.

La demostración se hace por inducción sobre la dimensión de V, n. Aquí veremos los casos $n = 1$ y $n = 2$ (que da la pauta para el razonamiento inductivo).

Si $n = 1$, cualquier base de V sirve. Sea $n = 2$ y supongamos que $q \neq 0$ (pues si $q = 0$ cualquier base sirve); tomemos un vector de $v_1 \in V$ tal que $q(v_1) \neq 0$ como primer vector de la base buscada. Los vectores conjugados con v_1 forman un subespacio de V de dimensión 1 (piense por qué: en general sería un subespacio de dimensión $n - 1$) en el cual podemos coger un vector v_2 para completar la base de vectores conjugados $\{v_1, v_2\}$.

Método de diagonalización por congruencia mediante operaciones elementales.- Para diagonalizar por congruencia una matriz simétrica, M, podemos emplear un procedimiento similar al que se describió hace unos capítulos para calcular la inversa de una matriz: colocamos I a la derecha de M, y vamos haciendo operaciones elementales en M para conseguir una matriz diagonal (haciendo ceros, pues); lo novedoso es que *hacemos las mismas operaciones en las filas que en las columnas.* De esa forma, en la matriz de la derecha, que empezó siendo I, las oe se realizan sólo en las filas. Al final, cuando a la izquierda tenemos una matriz diagonal, a la derecha nos queda una matriz, Q, que es la traspuesta de P: $(M|I)$ . ... $(D|Q = P^T)$. Veamos un ejemplo:

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \\ 3 & 5 & 8 \end{pmatrix}$$

escibimos I a su derecha:

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 5 & 0 & 1 & 0 \\ 3 & 5 & 8 & 0 & 0 & 1 \end{pmatrix}$$

a la segunda fila le restamos el doble de la primera, a la tercera le restamos el triple de la primera, y hacemos otro tanto con las columnas:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & -1 & -1 & -3 & 0 & 1 \end{pmatrix}$$

ahora restamos la segunda fila a la tercera (y lo mismo en las columnas:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{pmatrix}$$

La matriz P es la traspuesta de la que tenemos a la derecha:

$$P = \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Compruebe que, en efecto, $P^T \cdot M \cdot P$ es una matriz diagonal (con 1, -1, 0 en la diagonal).

Para comprender por qué ese procedimiento nos proporciona la base de vectores conjugados, recordemos que realizar operaciones elementales en las filas de una matriz M equivale a multiplicarla a su izquierda por una cierta matriz E (repase el capítulo 2) y realizarlas en las columnas es tanto como multiplicarla a la derecha por la matriz correspondiente; realizar las mismas operaciones elementales en las filas y en las columnas supone que las dos matrices son traspuesta una de la otra, de suerte que lo que hacemos es $E^T \cdot M \cdot E$; si en la matriz identidad realizamos las oe sólo en las filas, lo que resulta es $E^T \cdot I$, o sea, E^T. Así pues, si empezamos con el par $(M|I)$, acabamos con el $(E^T \cdot M \cdot E, E^T \cdot I)$ que será (D, E^T), La matriz $D = E^T \cdot M \cdot E$ es congruente con M, y por tanto representa a la misma forma cuadrática, q, en una base diferente. Si recordamos la fórmula de cambio de base para formas bilineales, que vimos hace poco, es evidente que E es la del cambio de la nueva base a la antigua.

Desde luego, podríamos haber obtenido directamente la matriz de cambio de base - y no su traspuesta - realizando las oe en las columnas de I en vez de realizarlas en las filas. Para ello, sería conveniente disponer la matriz I debajo de la matriz M, en lugar de colocarla a su derecha; sencillamente para no confundirnos en las operaciones. Por razones tipográficas, me resulta más sencillo hacerlo de la manera que he descrito.

Una vez diagonalizada una forma cuadrática, en la diagonal de su matriz aparecen los elementos $d_k = q(v_k)$, que pueden ser positivos, negativos o nulos. Podemos simplificar aún más esa matriz haciendo un pequeño retoque en la base: si tomamos $u_k = v_k/\sqrt{|d_k|}$ cuando $d_k \neq 0$, (y $u_k = v_k$ cuando $d_k = 0$), la matriz de q en la base $\{u_1, \ldots, u_n\}$ es una matriz diagonal que sólo tiene 1, -1 y 0.

En principio, la cantidad de términos positivos y negativos que aparecen al diagonalizar una forma cuadrática podría depender de qué base de vectores conjugados se elija. Sin embargo, eso no es así, como se demuestra fácilmente (aunque no lo haremos). El resultado se puede formular así:

Teorema de inercia.- Si M y N son dos matrices diagonales que representan a una misma forma cuadrática, q, en distintas bases, y si en la diagonal de M hay r términos positivos y s términos negativos, entonces en la diagonal de N también hay r términos positivos y s términos negativos.

Definición.- Al par (r, s) que recoge las cantidades de términos positivos y negativos en la diagonal se le llama *signatura de la forma cuadrática*. Llamamos signatura de una matriz cuadrada a la de la forma cuadrática sobre $\mathbf{R}^n$ a la que representa en la base canónica (o en cualquier otra base).

Nótese que $r + s$ es el rango de la forma cuadrática, de suerte que la signatura ya informa del rango.

Algunos libros definen signatura de maneras diferentes, por ejemplo, sólo como el número de términos positivos, o como la diferencia entre el número de términos positivos y el de negativos. Cuando se hace así, la signatura no contiene información suficiente para conocer el rango de la forma cuadrática.

La clasificación de las matrices que representan formas cuadráticas (o sea, la clasificación por congruencia) es ahora posible:

Proposición.- Dos formas cuadráticas difieren en un cambio de base si y sólo si tienen el mismo rango y la misma signatura. Del mismo modo, dos matrices son congruentes si y sólo si sus rangos coinciden y sus signaturas también.

Puesto que la signatura ya contiene al rango, el enunciado es redundante, pero lo formulo así para que coincida con los textos en que la signatura se define de otro modo. Conviene advertir que el resultado es válido para formas cuadráticas en espacios vectoriales sobre el cuerpo $\mathbf{R}$; sobre $\mathbf{C}$ el invariante que caracteriza esa equivalencia es el rango (pues en $\mathbf{C}$ no hay números positivos ni negativos), y sobre otros cuerpos la condición puede ser más complicada. De todos modos, nosotros consideramos solamente espacios vectoriales reales.

Entre las formas cuadráticas, tienen especial interés las que tienen un signo definido. El libro de Strang "Álgebra Lineal y sus aplicaciones"les dedica todo un capítulo (el 6º), cuya lectura es muy recomendable. El siguiente capítulo de estos apuntes también está consagrado al estudio de las formas positivas. No estará de más indicar aquí qué son y cómo se las puede reconocer.

Definición.- Una forma cuadrática, q, se dice *definida positiva* cuando $q(u) > 0, \forall u \neq 0$, y se dice *definida negativa* cuando $q(u) < 0, \forall u \neq 0$. Es *semidefinida positiva* cuando $q(u) \geq 0, \forall u$ y *semidefinida negativa* cuando $q(u) \leq 0, \forall u$. Finalmente, q es indefinida cuando toma valores positivos y negativos.

Se comprende fácilmente que si q es definida positiva, entonces $-q$ será definida negativa y viceversa. También es inmediato comprobar que:

Una CNS para que q sea definida positiva es que su signatura sea $(n, 0)$, siendo n la dimensión del espacio V.

Una CNS para que q sea definida negativa es que su signatura sea $(0, n)$.

Una CNS para que q sea semidefinida positiva es que su signatura sea $(r, 0)$.

Una CNS para que q sea semidefinida negativa es que su signatura sea $(0, s)$.

Para ver si una forma cuadrática es definida positiva no es preciso diagonalizarla. Hay un sencillo criterio (llamado *criterio de Sylvester*) que permite comprobarlo rápidamente. Lo enunciamos sin demostración:

Sea $M = (m_{ij})$ la matriz de la forma cuadrática q en una base, y sean $\Delta_1 = m_{11}$, $\Delta_2 = m_{11} \cdot m_{22} - m_{12} \cdot m_{21}$, ..., $\Delta_n = |M|$ los menores angulares de la matriz M. Entonces una condición necesaria y suficiente para que q sea definida positiva es que todos esos menores angulares sean positivos: $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_n > 0$.

Observación.- Una vía alternativa para discutir este punto es calcular los autovalores de la matriz M: si todos ellos son positivos, entonces q es definida positiva, y no lo será en caso contrario. Es de advertir que los autovalores no se conservan al cambiar M por $P^T \cdot M \cdot P$, pero sí se mantiene el signo.

Puede verse la demostración de estos criterios de positividad en el libro de Strang antes citado (páginas 279 y siguientes).

Comentario sobre la utilidad de las formas cuadráticas en el estudio de máximos y mínimos.- Así como la discusión de la naturaleza de los puntos críticos de las funciones de una variable se basa en el signo que tenga la segunda derivada, una vez que la primera se anula (si f'' es positiva, estamos ante un mínimo de f; si es negativa, ante un máximo), en las funciones de varias variables hay que prestar atención a la matriz hessiana (que contiene las derivadas parciales de segundo orden), que es simétrica y representa una forma cuadrática cuya signatura contiene mucha información sobre el punto crítico.

Para verlo con claridad, pensemos que en el caso de una variable, el polinomio de Taylor $P_2(x) = f(a) + \frac{f''(a)}{2}(x-a)^2$ constituye una aproximación óptima a la función f cerca del punto a; la gráfica de ese polinomio es una parábola, que tiene un máximo o un mínimo en a según sea el signo de $f''(a)$.

En el caso de una función de dos variables, la gráfica $z = f(x, y)$ es una superficie. En un punto $P = (a, b)$ en el que se anulen las dos derivadas parciales $f_x(P) = f_y(P) = 0$, el plano tangente es horizontal, y la aproximación dada por el segundo polinomio de Taylor $f(P) + \frac{1}{2}[f_{xx}(P)(x-a)^2 + 2f_{xy}(P)(x-a)(y-b) + f_{yy}(P)(y-b)^2]$ es la que nos habla de la naturaleza del punto crítico: si el término de segundo grado es siempre positivo (o sea, si la forma cuadrática $f_{xx}(P)(x-a)^2 + 2f_{xy}(P)(x-a)(y-b) + f_{yy}(P)(y-b)^2$ es definida positiva), estamos ante un mínimo de f; si ese término es siempre negativo, ante un máximo; y si es indefinida, nos hallamos en un punto de silla. La gráfica de ese polinomio de Taylor es una superficie conocida como paraboloide; puede tener la forma de un cuenco vuelto hacia arriba (lo que nos da un mínimo), hacia abajo (correspondiente a un máximo) o de una silla de montar (para un punto que no es máximo ni mínimo).

Se comprende la importancia del estudio de la forma cuadrática definida por la matriz hessiana a la hora de buscar máximos y mínimos en funciones de dos variables (o de más). Por otra parte, disponemos de herramientas variadas para discutir el signo de esa forma cuadrática (diagonalización por oe, autovalores), pero los libros de Cálculo suelen decantarse por el criterio de Sylvester. Nada nos obliga a renunciar a otros métodos, si lo preferimos.

Ejemplo.- La función $f(x, y) = xy + \frac{1}{x} - \frac{1}{y}$ tiene por derivadas parciales $f_x = y - \frac{1}{x^2}$, $f_y = x + \frac{1}{y^2}$, que se anulan en el punto $P = (-1, 1)$. Las segundas derivadas, $f_{xx} = \frac{2}{x^3}$, $f_{xy} = 1$, $f_{yy} = \frac{-2}{y^3}$ toman en ese punto los valores $-2, 1, -2$ respectivamente, por lo que la matriz hessiana en el punto crítico es $H = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ que corresponde a una forma cuadrática definida negativa, puesto que sus autovalores son -1 y -3. En el punto $(-1, 1)$, la función f tiene un máximo.

Consideraciones finales.- Aunque el estudio se ha ceñido a las formas bilineales y cuadráticas en espacios vectoriales reales, también tiene interés considerar el caso de los espacios complejos (por ejemplo, en la Mecánica Cuántica). Esa tarea no corresponde a este nivel, pero no me resisto a escribir aquí que hay que empezar por cambiar la propia definición, y sustituir la condición $f(u, v) = f(v, u)$ por $f(u, v) = \overline{f(v, u)}$, donde la línea representa la conjugación compleja. Se habla entonces de formas sesquilineales (en lugar de bilineales) y hermíticas (en vez de simétricas). El alumno interesado puede explorar este terreno en múltiples textos y en la red.

Ejercicios

1.- Si S es un subespacio vectorial de V, discuta si ha de darse la igualdad $S^{00} = S$ o no, y qué relación hay entre ellos.

2.- Demuestre que si una forma cuadrática nunca toma el valor 0 (salvo para el vector nulo, se entiende), entonces es definida positiva o definida negativa. (Sugerencia: diagonalice la forma cuadrática; si el rango no es máximo, ya es fácil; si el rango es máximo y en la diagonal hay términos de ambos signos, quiere decir que hay dos vectores conjugados en la base tales que $q(u) > 0$ y $q(v) < 0$. Sea $f(t) = q(tu + (1-t)v)$; entonces $f(0)$ es positivo y $f(1)$ negativo, por lo que f tomará el valor 0 alguna vez). (Para otra estrategia, diagonalice la forma cuadrática).

3.- Clasifique la forma cuadrática $q(x, y, z) = y^2 + z^2 + 2xy + 2xz + 4yz$ y encuentre una base de vectores conjugados.

4.- Clasifique las formas cuadráticas cuyas matrices en la base canónica de $\mathbf{R}^4$ son:

$$\begin{pmatrix} 1 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 1 & 0 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 & -1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}$$

5.- Discuta si el conjunto de los vectores autoconjugados para una forma cuadrática en V es o no un subespacio vectorial del espacio V.

6.- Demuestre que una forma definida (positiva o negativa) es no degenerada, pero el recíproco no es cierto. Del mismo modo, si la forma cuadrática q es definida, ningún vector es autoconjugado (salvo el nulo). Discuta si es o no cierta la afirmación recíproca.