PARTE III FORMAS BILINEALES

CAPÍTULO 11 FORMAS BILINEALES Y CUADRÁTICAS

§ 90. Propiedades generales de las formas bilineales y cuadráticas

Veamos las funciones numéricas $\varphi(x, y)$ de dos argumentos vectoriales x, y de cierto espacio lineal K_n, dado sobre un campo numérico P; x, y toman los valores de P. Una función $\varphi(x, y)$ se llama *forma bilineal*, si para cualesquiera dos vectores x, y, $z \in K_n$ y todo número $\alpha \in P$ se verifican las correlaciones

$$\varphi(x+z, y) = \varphi(x, y) + \varphi(z, y), \quad \varphi(\alpha x, y) = \alpha\varphi(x, y),$$
$$\varphi(x, y+z) = \varphi(x, y) + \varphi(x+z), \quad \varphi(x, \alpha y) = \alpha\varphi(x, y). \quad (90.1)$$

Las primeras dos correlaciones de (90.1) significan la linealidad de la forma $\varphi(x, y)$ respecto del primer argumento, las dos últimas, la linealidad respecto del segundo argumento.

Es fácil comprobar que la suma de dos formas bilineales, como también el producto de una forma bilineal por un número será nuevamente una forma bilineal. Por esta razón, el conjunto de todas las formas bilineales, definidas sobre un mismo espacio K_n, que toman los valores de un mismo campo numérico P será un espacio lineal. En este caso el «cero» del espacio dado será la forma bilineal $0(x, y)$, para la cual $0(x, y) = 0$, cualesquiera que sean x, y. $0(x, y)$ se denomina forma bilineal *nula*.

Anteriormente ya nos encontramos con una función de tal índole. Comparando (27.1) y (90.1), se nota con facilidad que un producto escalar en un espacio euclídeo es una forma bilineal. Recordando el papel importante que desempeñó el producto escalar al estudiar los espacios euclídeos y los operadores lineales que actúan en los mismos, podemos suponer que el estudio de las formas bilineales será también útil.

Entre las formas bilineales las simétricas y las antisimétricas atraen un interés singular. Una forma bilineal $\varphi(x, y)$ se llama *simétrica*, si para cualesquiera vectores $x, y \in K_n$ se verifica la desigualdad

$$\varphi(x, y) = \varphi(y, x).$$

En cambio, si para cualesquiera $x, y \in K_n$

$$\varphi(x, y) = -\varphi(y, x),$$

la forma bilineal se denomina *antisimétrica*.

Toda forma bilineal antisimétrica $\varphi(x, y)$ se anula, si los argumentos coinciden. En efecto, como $\varphi(x, x) = -\varphi(x, x)$, entonces $\varphi(x, x) = 0$. Algo sorprendente parece el otro hecho, vinculado con los valores de una forma bilineal simétrica al coincidir los argumentos. A saber, cada forma bilineal simétrica $\varphi(x, y)$ se define unívocamente por sus valores, siendo coincidentes los argumentos. Efectivamente, sean x, y cualesquiera vectores de K_n. Tomando en consideración la simetría de la forma $\varphi(x, y)$, tenemos

$$\varphi(x + y, x + y) = \varphi(x, x) + \varphi(y, y) + 2\varphi(x, y), \qquad (90.2)$$

de donde se desprende que

$$\varphi(x, y) = \frac{1}{2}\{\varphi(x + y, x + y) - \varphi(x, x) - \varphi(y, y)\}. \qquad (90.3)$$

La fórmula obtenida demuestra que la afirmación enunciada es justa, puesto que el segundo miembro de la correlación es una forma bilineal simétrica.

Una forma bilineal se descompone unívocamente en la suma de las formas bilineales simétrica y antisimétrica. Esta descomposición puede escribirse explícitamente

$$\varphi(x, y) = \frac{1}{2}\{\varphi(x, y) + \varphi(y, x)\} + \frac{1}{2}\{\varphi(x, y) - \varphi(y, x)\}. \qquad (90.4)$$

Es fácil comprobar que los primeros dos sumandos en el miembro derecho dan una forma bilineal simétrica y los dos últimos, antisimétrica. Si se admite la existencia de alguna otra descomposición, entonces, al sustituir los argumentos iguales, habremos de hacer una deducción sobre la univocidad de la definición de la parte simétrica de la descomposición y, consecuentemente, de la descomposición en total.

Si una forma bilineal no es simétrica, en lugar de (90.2) tendremos

$$\varphi(x + y, x + y) = \varphi(x, x) + \varphi(y, y) + \varphi(x, y) + \varphi(y, x).$$

Por consiguiente,

$$\frac{1}{2}\{\varphi(x, y) + \varphi(y, x)\} = \frac{1}{2}\{\varphi(x + y, x + y) - \varphi(x, x) - \varphi(y, y)\}. \qquad (90.5)$$

Comparando la correlación obtenida con (90.3), concluimos que para la forma bilineal antisimétrica su parte simétrica se define unívocamente por los valores de la forma, siendo coincidentes los argumentos.

A la par con las formas bilineales consideraremos también las así llamadas formas cuadráticas. Sea $\varphi(x, y)$ una forma bilineal en el espacio K_n. Se denomina *forma cuadrática* la función numérica $\varphi(x, x)$ de un solo argumento vectorial $x \in K_n$, la cual se obtiene de la forma bilineal $\varphi(x, y)$ sustituyendo el vector y por el vector x.

En general, partiendo de la forma cuadrática, no se puede restablecer unívocamente la forma bilineal que la ha engendrado. Pero, según se deduce de la fórmula (90.3), existe una forma bilineal simétrica y sólo una de la cual puede ser obtenida la forma cuadrática inicial. Esta forma bilineal se llama *polar* respecto a la forma cuadrática dada. El conjunto de todas las formas bilineales que engendran una misma forma cuadrática puede ser obtenido por sumación de la forma bilineal polar con una forma antisimétrica arbitraria. Es por eso que al utilizar formas bilineales para estudiar las propiedades de las formas cuadráticas resulta suficiente limitarse sólo a la consideración de las formas bilineales simétricas.

El hecho de que una forma bilineal no puede ser restablecida según la cuadrática se debe a que la última no proporciona ninguna información sobre la parte antisimétrica, cualquiera que sea la forma bilineal.

LEMA 90.1 *Las formas bilineales antisimétricas, y sólo ellas, se anulan, cuando todos los argumentos coinciden.*

DEMOSTRACIÓN. Ya hemos señalado que si $\varphi(x, y)$ es una forma antisimétrica, entonces $\varphi(x, x) = 0$ para cualquier x. En cambio, si $\varphi(x, x) = 0$ para todo x, entonces de la correlación (90.5) se infiere que para todos los vectores x, y se verifica la igualdad $\varphi(x, y) + \varphi(y, x) = 0$, es decir, la forma bilineal $\varphi(x, y)$ es antisimétrica.

La comparación de las propiedades que poseen el producto escalar y las correlaciones (90.1) muestra que en un espacio unitario el producto escalar no es, estrictamente, una forma bilineal. En un espacio complejo las formas bilineales hermitianas están estrechamente vinculadas con el producto escalar. La función numérica $\varphi(x, y)$ se llama forma bilineal *hermitiana*, si para cualesquiera vectores x, y, $z \in K_n$ y todo número α del campo de números complejos P se verifican las correlaciones

$$\varphi(x+z, y) = \varphi(x, y) + \varphi(z, y), \qquad \varphi(\alpha x, y) = \alpha\varphi(x, y),$$
$$\varphi(x, y+z) = \varphi(x, y) + \varphi(x, z), \qquad \varphi(x, \alpha y) = \bar{\alpha}\varphi(x, y).$$

La raya significa aquí una conjugación compleja.

En este caso también la suma de dos formas bilineales hermitianas y asimismo el producto de una forma bilineal hermitiana por un número será una forma bilineal hermitiana. Por esto el conjunto de todas las formas bilineales hermitianas que están definidas sobre un espacio complejo y que toman valores complejos es un espacio lineal complejo.

La forma bilineal hermitiana se denomina *simétrica hermitiana*, si para cualesquiera vectores x, $y \in K_n$ se tiene

$$\varphi(x, y) = \overline{\varphi(y, x)}.$$

Si para cualesquiera x, $y \in K_n$

$$\varphi(x, y) = -\overline{\varphi(y, x)},$$

la forma se llama *antisimétrica hermitiana*. En los vectores coincidentes la forma antisimétrica hermitiana toma valores imaginarios puros, mientras que la forma simétrica hermitiana, valores reales. Ahora, toda forma bilineal hermitiana se define unívocamente por sus valores si los argumentos son coincidentes. Pero, en vez de (90.3), es lícita la correlación

$$\varphi(x, y) = \frac{1}{4}\{\varphi(x+y, x+y) - \varphi(x-y, x-y) + \\ + i\varphi(x+iy, x+iy) - i\varphi(x-iy, x-iy)\}. \quad (90.6)$$

De esta correlación se infiere, en particular, que

Entre las formas bilineales hermitianas la forma nula, y sólo ella, toma los valores nulos cuando todos los argumentos coinciden.

En este caso también la forma bilineal hermitiana puede ser unívocamente representada como una suma de la forma simétrica hermitiana y la antisimétrica hermitiana, con la particularidad de que

$$\varphi(x, y) = \frac{1}{2}\{\varphi(x, y) + \overline{\varphi(y, x)}\} + \frac{1}{2}\{\varphi(x, y) - \overline{\varphi(y, x)}\}. \quad (90.7)$$

Las demostraciones de los hechos enunciados para las formas hermitianas casi no se diferencian de las demostraciones correspondientes para las formas bilineales.

Se llama forma cuadrática *hermitiana* la función numérica $\varphi(x, x)$ de un solo argumento vectorial $x \in K_n$, la cual se obtiene de la función bilineal hermitiana $\varphi(x, y)$ al sustituir el vector y por el vector x. A diferencia de las formas cuadráticas, a base de una forma cuadrática hermitiana se restablece unívocamente la forma bilineal hermitiana que engendra la citada forma cuadrática hermitiana. Este restablecimiento se realiza de acuerdo con la fórmula (90.6) y la correspondiente forma bilineal se llama también *polar* respecto a la forma cuadrática de partida.

La posibilidad de restablecer unívocamente una forma bilineal hermitiana sobre la base de la forma cuadrática hermitiana engendrada por la primera se explica por una estrecha relación que existe entre las formas bilineales simétricas hermitianas y antisimétricas hermitianas.

LEMA 90.2 *Si $\varphi(x, y)$ es una forma bilineal simétrica (antisimétrica) hermitiana, entonces $\psi(x, y) = i\varphi(x, y)$ será una forma bilineal antisimétrica (simétrica) hermitiana.*

DEMOSTRACION. Sea, por ejemplo, $\varphi(x, y)$ una forma simétrica hermitiana. Entonces, para todos los vectores x, y se tiene

$$\psi(x, y) = i\varphi(x, y) = \varphi(ix, y) = \overline{\varphi(y, ix)} = -\overline{i\varphi(y, x)} = -\overline{\psi(y, x)},$$

es decir, $\psi(x, y)$ es una forma antisimétrica hermitiana. El caso de la forma antisimétrica hermitiana $\varphi(x, y)$ se considera de modo análogo.

En lo que sigue trataremos frecuentemente las formas cuadráticas hermitianas engendradas por las formas bilineales simétricas hermitianas.

LEMA 90.3 *Entre las formas bilineales hermitianas las simétricas, y sólo ellas, engendran formas cuadráticas hermitianas reales.*

DEMOSTRACION. Anteriormente ya se ha observado que las formas simétricas hermitianas toman valores reales cuando coinciden los argumentos. Supondremos ahora que una forma cuadrática hermitiana $\varphi(x, x)$ toma sólo los valores reales. Conforme a (90.6), para la forma bilineal polar $\varphi(x, y)$ tenemos

$$\begin{aligned}\varphi(y, x) &= \frac{1}{4}\{\varphi(y+x, y+x) - \varphi(y-x, y-x) + \\ &\qquad + i\varphi(y+ix, y+ix) - i\varphi(y-ix, y-ix)\} = \\ &= \frac{1}{4}\{\varphi(x+y, x+y) - \varphi(x-y, x-y) + i\varphi(x-iy, x-iy) - \\ &\qquad - i\varphi(x+iy, x+iy)\} = \frac{1}{4}\{\overline{\varphi(x+y, x+y)} - \overline{\varphi(x-y, x-y)} + \\ &\qquad + \overline{i\varphi(x+iy, x+iy)} - \overline{i\varphi(x-iy, x-iy)}\} = \overline{\varphi(x, y)}.\end{aligned}$$

COROLARIO. *Entre las formas bilineales hermitianas las antisimétricas, y sólo ellas, engendran formas cuadráticas hermitianas imaginarias puras.*

COROLARIO. *Ninguna forma bilineal antisimétrica hermitiana puede engendrar una forma cuadrática hermitiana real.*

Según se deduce de las propiedades de linealidad de las formas bilineales y bilineales hermitianas respecto de cada argumento, $\varphi(0, 0) = 0$ para cualquier forma cuadrática $\varphi(x, x)$. Sin embargo, en el caso general pueden existir también los vectores no nulos x, para los cuales $\varphi(x, x) = 0$. Tales vectores se llamarán *isótropos*. El concepto de isotropía está ligado sólo con la forma cuadrática. Por ello, unos vectores, isótropos para una forma cuadrática, pueden no serlo para otra forma cuadrática, y viceversa. En particular, el lema 90.1 significa que para una forma cuadrática engendrada por una forma bilineal antisimétrica, todos los vectores del espacio K_n, salvo el nulo, son isótropos.

De las formas reales hermitianas y ordinarias son de mayor uso aquellas que toman los valores de un mismo signo para todos los argumentos vectoriales. La forma cuadrática real $\varphi(x, x)$ se denomina *definida positiva*, si $\varphi(x, x) > 0$ para todo $x \neq 0$. La forma se llama *no negativa*, si para todo $x \neq 0$ se cumple la desigualdad $\varphi(x, x) \geqslant 0$. Análogamente se definen las formas cuadráticas *no positivas* y *definidas negativas*.

Como regla, solamente las formas cuadráticas definidas positivas y definidas negativas se llaman *de signo constante*. Pero, a veces, así se llaman también las formas cuadráticas no negativas y no positivas. Con el fin de evitar toda clase de equivocaciones, las formas cuadráticas definidas positivas y definidas negativas se llamarán, cuando sea necesario, *estrictamente* de signo constante.

Si una forma cuadrática real es de signo constante, la forma bilineal hermitiana u ordinaria que la engendra se llamará también definida positiva, no negativa, etc.

Si una forma cuadrática real $\varphi(x, x)$ es estrictamente de signo constante, no tiene vectores isótropos. En el caso de las formas bilineales reales y bilineales simétricas hermitianas $\varphi(x, y)$, las correspondientes formas cuadráticas serán reales y para ellas resulta cierta la afirmación recíproca. A saber, tiene lugar el

TEOREMA 90.1. *Supongamos que una forma cuadrática $\varphi(x, x)$ está engendrada por la forma bilineal real o bilineal simétrica hermitiana $\varphi(x, y)$. Si $\varphi(x, x)$ no tiene vectores isótropos, es estrictamente de signo constante.*

DEMOSTRACIÓN. Como ya se ha indicado, la forma cuadrática $\varphi(x, x)$ es real. En ambos casos toma en vectores colineales los valores de un mismo signo. Supongamos que $\varphi(x, x)$ no es estrictamente de signo constante. En este caso existen tales vectores linealmente independientes u, v que $\varphi(u, u) > 0$ y $\varphi(v, v) < 0$. Para todo número real α

$$\varphi(u + \alpha v, u + \alpha v) = \varphi(u, u) + \alpha(\varphi(u, v) + \varphi(v, u)) + \alpha^2\varphi(v, v). \qquad (90.8)$$

El segundo miembro de esta igualdad es un polinomio de segundo grado respecto de α. Sus coeficientes son reales, lo que se determina por el carácter real de la forma cuadrática $\varphi(x, x)$ y el lema 90.3. Puesto que $\varphi(u, u)$ y $\varphi(v, v)$ son de signos opuestos, el polinomio (90.8) tendrá dos raíces reales. Sea α_0 una de ellas. Esto significa que $\varphi(u + \alpha_0 v, u + \alpha_0 v) = 0$. Sin embargo, el vector $u + \alpha_0 v$ es no nulo por ser los vectores u, v linealmente independientes, por lo cual, la anulación en él de la forma cuadrática no es posible por hipótesis del teorema. La contradicción obtenida da por terminado la demostración del mismo.

No es casual que en el teorema 90.1 nos limitamos a la consideración de las formas cuadráticas engendradas sólo por las formas bilineal real y bilineal simétrica hermitiana. Ninguna otra forma bilineal puede conducir a una forma cuadrática real. Nos queda considerar, de hecho, sólo una forma bilineal en un espacio complejo. Pero tal forma bilineal no puede engendrar una forma cuadrática real que no sea idénticamente igual a cero. Si para cierto vector u una forma cuadrática toma el valor real $\varphi(u, u)$ que no es igual a cero, entonces $\varphi(\alpha u, \alpha u) = \alpha^2\varphi(u, u)$ será un número complejo, cualquiera que sea α complejo con las partes imaginaria pura y real no nulas. Así pues,

Para las formas cuadráticas reales la condición necesaria y suficiente de que estas formas no tengan vectores isótropos consiste en el mantenimiento estricto de signo constante.

La forma bilineal compleja genera siempre una forma cuadrática, que tiene vectores isótropos, siempre que esté definida en un espacio lineal cuya dimensión es superior a uno. En efecto, si suponemos que esto no es así, se hallarán siempre unos vectores linealmente independientes u, v, para los cuales $\varphi(u, u) \neq 0$, $\varphi(v, v) \neq \neq 0$. Pero, de conformidad con (90.8), el vector $u + \alpha v$ será isótropo, elegido de manera adecuada el número complejo α. La forma compleja bilineal hermitiana puede generar una forma cuadrática que no tiene vectores isótropos. Según se deduce de nuestras investigaciones,

Para que una forma cuadrática engendrada por la forma bilineal hermitiana no tenga vectores isótropos, es suficiente que la parte real (o imaginaria) de la forma cuadrática sea estrictamente de signo constante.

Ejercicios.

1. Demuéstrese que para toda forma bilineal $\varphi(x, y)$ se verifican las desigualdades $\varphi(0, y) = \varphi(x, 0) = 0$, cualesquiera que sean x, $y \in K_n$.
2. Determínense la dimensión y la base de un espacio lineal de formas bilineales.
3. Demuéstrese que los conjuntos de formas bilineales simétricas y antisimétricas forman subespacios en un espacio lineal de todas las formas bilineales.
4. Demuéstrese que un espacio de todas las formas bilineales es la suma directa de los subespacios de las formas bilineales simétricas y antisimétricas.
5. Demuéstrese que un conjunto de todas las formas cuadráticas forman un espacio lineal. Determínense su dimensión y la base.
6. ¿Formarán los subespacios lineales los siguientes conjuntos de formas cuadráticas:

formas cuadráticas de signo constante,
formas cuadráticas que toman los valores reales,
formas cuadráticas que no tienen vectores isótropos,
formas cuadráticas, para las cuales todos los vectores del conjunto dado son isótropos?

7. Demuéstrese que para toda forma cuadrática, dada en un espacio normalizado, existe tal número α que para todo x

$$|\varphi(x, x)| \leqslant \alpha \|x\|^2.$$

8. Supongamos que la forma cuadrática $\varphi(x, x)$ es estrictamente de signo constante y la forma cuadrática $\psi(x, x)$ es arbitraria. Demuéstrese que exista tal número β, que para todo x

$$|\psi(x, x)| \leqslant \beta\varphi(x, x).$$

9. Demuéstrese que una forma cuadrática no es estrictamente de signo constante cuando, y sólo cuando, el conjunto de vectores isótropos y el vector nulo forman un subespacio lineal.

10. Examínense los ejercicios 1—9 para las formas bilineales hermitianas y las cuadráticas. ¿Serán verídicas todas las afirmaciones enunciadas?

11. Supongamos que en el espacio complejo K_n un subespacio L sólo consiste de los vectores isótropos de la forma bilineal hermitiana $\varphi(x, y)$ y el vector nulo. Demuéstrese que $\varphi(u, v) = 0$ para cualesquiera vectores $u, v \in L$.

§ 91. Matrices de las formas bilineales y cuadráticas

Investiguemos una forma bilineal $\varphi(x, y)$ definida en el espacio K_n. Elijamos en K_n dos bases fijadas $e_1, e_2, \ldots, e_n$ y $q_1, q_2, \ldots, q_n$ y sea

$$x = \sum_{i=1}^{n} \xi_i e_i, \qquad y = \sum_{j=1}^{n} \eta_j q_j.$$

En virtud de las propiedades (90.1) tenemos

$$\varphi(x, y) = \varphi\left(\sum_{i=1}^{n} \xi_i e_i, \quad \sum_{j=1}^{n} \eta_j q_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi(e_i, q_j)\, \xi_i \eta_j. \quad (91.1)$$

Designemos, como antes, con x_e e y_q las matrices de dimensiones $n \times 1$, compuestas por las coordenadas de los vectores x y y en las bases correspondientes, y mediante G_{eq}, la matriz de orden n con los elementos $g_{ij}^{(eq)} = \varphi(e_i, q_j)$. La correlación (91.1) significa que

$$\varphi(x, y) = x_e' G_{eq} y_q \qquad (91.2)$$

De este modo, siendo fijadas las bases en el espacio K_n, la forma bilineal puede ser representada en la forma matricial (91.2).

G_{eq} se llama *matriz de la forma bilineal* y con las bases fijadas se define unívocamente. Si suponemos que para la forma $\varphi(x, y)$ existe, además de (91.2), otra representación análoga con cierta matriz F_{eq}, entonces, al poner $x = e_i$, $y = q_j$ obtenemos en seguida que $f_{ij}^{(eq)} = \varphi(e_i, q_j)$, es decir, $F_{eq} = G_{eq}$.

Cabe señalar que el segundo miembro de (91.2) define, para cualquier matriz G_{eq}, cierta forma bilineal. El cumplimiento de las correlaciones (90.1) se infiere directamente de las propiedades correspondientes de las operaciones matriciales. Así se establece, con las bases fijadas en K_n, una correspondencia biunívoca entre las formas bilineales y las matrices cuadradas.

Al cambiar las bases en K_n, la matriz de la forma bilineal, desde luego, varía. Sea P una matriz de la transformación de coordenadas al pasar de la base $e_1, e_2, \ldots, e_n$ a la $f_1, f_2, \ldots, f_n$ y sea Q la matriz de la transformación de coordenadas al pasar de $q_1, q_2, \ldots, q_n$ a $t_1, t_2, \ldots, t_n$. De acuerdo con (63.3)

$$x_e = Px_f, \qquad y_q = Qy_t, \tag{91.3}$$

por lo cual de (91.2) se desprende que

$$\varphi(x, y) = x_e' G_{eq} y_q = x_f' P' G_{eq} Q y_t.$$

Pero, por otra parte,

$$\varphi(x, y) = x_f' G_{ft} y_t.$$

Por consiguiente,

$$G_{ft} = P' G_{eq} Q. \tag{91.4}$$

Como las matrices P y Q son regulares, entonces, de acuerdo con la terminología introducida en el § 64, las matrices G_{ft} y G_{eq} se llamarán equivalentes. Según se ha señalado anteriormente, las matrices equivalentes de un mismo orden, y sólo ellas, tienen rangos iguales. Esto es testimonio de que el rango de una matriz de la forma bilineal no depende de las bases elegidas y es una característica de la misma forma. Llamémoslo *rango* de la forma bilineal. Una forma bilineal se llamará *regular*, si es regular su matriz. Como característica de una forma bilineal sirve también la diferencia entre la dimensión del espacio K_n y el rango de la forma. Se llamará dicha característica *defecto* de la forma bilineal.

De los resultados del § 64 se desprende que todas las matrices de un mismo rango son equivalentes a la matriz diagonal con los elementos 0 y 1. En el lenguaje de las formas bilineales este hecho atestigua que para una forma arbitraria de rango r siempre pueden indicarse tales bases $f_1, f_2, \ldots, f_n$ y $t_1, t_2, \ldots, t_n$, en las cuales la forma tendrá una expresión más simple. A saber, si

$$x = \sum_{i=1}^{n} \tau_i f_i, \qquad y = \sum_{j=1}^{n} \nu_j t_j,$$

entonces

$$\varphi(x, y) = \sum_{i=1}^{r} \tau_i \nu_i.$$

La elección separada de las bases para toda forma bilineal variable se realiza raramente. Con mucha más frecuencia se utiliza una base común. Sea $e_1, e_2, \ldots, e_n$ cierta base de K_n y

$$x = \sum_{i=1}^{n} \xi_i e_i, \qquad y = \sum_{j=1}^{n} \gamma_j e_j.$$

En este caso, por analogía con (91.1), obtenemos la siguiente representación de la forma bilineal:

$$\varphi(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi(e_i, e_j)\, \xi_i \gamma_j,$$

o, en la escritura matricial,

$$\varphi(x, y) = x'_e G_e y_e. \tag{91.5}$$

Aquí G_e es una matriz con los elementos $g_{ij}^{(e)} = \varphi(e_i e_j)$. En lo sucesivo la matriz G_e se llamará siempre matriz de la forma bilineal. Si, de nuevo, P es una matriz de la transformación de coordenadas al pasar de la base $e_1, e_2, \ldots, e_n$ a la $f_1, f_2, \ldots, f_n$, entonces las matrices G_e y G_f de la misma forma bilineal $\varphi(x, y)$ estarán ligadas entre sí, de acuerdo con (91.4), por la correlación

$$G_f = P' G_e P. \tag{91.6}$$

Las matrices G_e y G_f, ligadas mediante la correlación (91.6), siendo regular la matriz P, se denominan *congruentes*. Las matrices congruentes son siempre equivalentes. Lo recíproco, desde luego, en el caso general no es cierto.

Todo lo dicho acerca de las formas bilineales puede ser aplicado con cambios insignificantes a las formas bilineales hermitianas. Toda forma hermitiana se representa unívocamente en forma matricial

$$\varphi(x, y) = x'_e G_{eq} \bar{y}_q,$$

siendo fijadas las bases $e_1, e_2, \ldots, e_n$ y $q_1, q_2, \ldots, q_n$. Al pasar a las otras bases $f_1, f_2, \ldots, f_n$ y $t_1, t_2, \ldots, t_n$, en lugar de (91.4) tendremos

$$G_{ft} = P' G_{eq} \bar{Q}.$$

Si los argumentos de la forma bilineal hermitiana están dados en una misma base, la anotación matricial de la forma es análoga a (91.6). A saber

$$\varphi(x, y) = x'_e G_e \bar{y}_e \tag{91.7}$$

Al pasar a la nueva base, las matrices de la forma quedarán ligadas entre sí por la correlación

$$G_f = P' G_e \bar{P}$$

y se dirá que estas matrices son *congruentes según Hermite*.

Ahora se puede establecer la relación entre la expresión de una forma bilineal y la de su matriz. Si la forma es simétrica, para cualquier base $e_1, e_2, \ldots, e_n$ se tiene

$$g_{ij}^{(e)} = \varphi(e_i, e_j) = \varphi(e_j, e_i) = g_{ji}^{(e)},$$

donde $G_e = G_e'$ y la matriz G_e de la forma $\varphi(x, y)$ es *simétrica*. En cambio, si la forma es antisimétrica, entonces

$$g_{ij}^{(e)} = \varphi(e_i, e_j) = -\varphi(e_j, e_i) = -g_{ji}^{(e)},$$

es decir, $G_e = -G_e'$. En este caso la matriz G_e se llama también *antisimétrica*.

La afirmación recíproca es asimismo cierta. Si en una base la matriz de la forma es simétrica (antisimétrica), la forma bilineal que la genera será también simétrica (antisimétrica). Sea $G_e = G_e'$, entonces

$$\varphi(y, x) = y_e'G_ex_e = (y_e'G_ex_e)' = x_e'G_e'y_e = x_e'G_ey_e = \varphi(x, y).$$

Si, en cambio, $G_e' = -G_e$, entonces

$$\varphi(x, y) = y_e'G_ex_e = (y_e'G_ex_e)' = x_e'G_e'y_e = -x_e'G_ey_e = -\varphi(x, y).$$

Las afirmaciones análogas subsisten también respecto de la relación existente entre la forma bilineal hermitiana y su matriz. Si la forma hermitiana es simétrica, entonces

$$g_{ij}^{(e)} = \varphi(e_i, e_j) = \overline{\varphi(e_j, e_i)} = \overline{g_{ji}^{(e)}},$$

es decir, $G_e = G_e^*$ y la matriz G_e de la forma $\varphi(x, y)$ es *hermitiana*. Si la forma es antisimétrica hermitiana, entonces

$$g_{ij}^{(e)} = \varphi(e_i, e_j) = -\overline{\varphi(e_j, e_i)} = -\overline{g_{ji}^{(e)}},$$

donde $G_e = -G_e^*$. En este caso la matriz G_e se denomina *antihermitiana*.

Son ciertas también las afirmaciones recíprocas. Sea $G_e = G_e^*$, entonces para una forma bilineal hermitiana generadora tenemos

$$\varphi(y, x) = y_e'G_e\overline{x}_e = (y_e'G_e\overline{x}_e)' = \overline{x}_e'G_e'y_e = \\ = \overline{x_e'G_e^*\overline{y}_e} = \overline{x_e'G_e\overline{y}_e} = \overline{\varphi(x, y)}.$$

Para el caso en que $G_e = -G_e^*$ encontramos

$$\varphi(y, x) = y_e'G_e\overline{x}_e = (y_e'G_e\overline{x}_e)' = \overline{x}_e'G_e'y_e = \\ = \overline{x_e'G_e^*\overline{y}_e} = -\overline{x_e'G_e\overline{y}_e} = -\overline{\varphi(x, y)}.$$

La matriz de la forma bilineal nula sólo se compone de elementos nulos, es decir, es una matriz nula. Esta es la única matriz que al mismo tiempo es simétrica y antisimétrica, al igual que la forma nula.

Ya se ha notado que existe una relación estrecha entre las formas bilineales simétricas y las cuadráticas. Esta relación se observa claramente en el nivel matricial. Para una forma bilineal $\varphi(x, y)$ es válida la correlación matricial (91.5). Para la forma cuadrática correspondiente tenemos

$$\varphi(x, x) = x_e' G_e x_e. \tag{91.8}$$

Siendo fijada la base $e_1, e_2, \ldots, e_n$, cada anotación del tipo (91.8) define, para cualquier matriz G_e, cierta forma cuadrática. La matriz G_e en (91.8) ya no se llama matriz de la forma bilineal, sino *matriz de la forma cuadrática.*

Si para las formas bilineales existe una correspondencia biunívoca entre las formas y las matrices de éstas, siendo fijada una base en K_n, en las circunstancias que se describen tal correspondencia ya no existe. Cada forma cuadrática puede ser definida por un conjunto de sus matrices. Dicho conjunto contiene una sola matriz simétrica y la diferencia entre cualesquiera dos matrices del conjunto dado es una matriz antisimétrica.

De este modo, cualquier forma cuadrática ordinaria puede ser definida siempre por una matriz simétrica. Al pasar a otra base, las matrices de la forma cuadrática varían de conformidad con (91.6). Por esta razón concluimos otra vez que los problemas de investigación de las formas bilineales simétricas y de las cuadráticas están estrechamente entrelazados. Para las formas cuadráticas hermitianas ya no es así, puesto que entre éstas y las formas bilineales hermitianas existe una correspondencia biunívoca y la misma correspondencia se mantiene entre sus matrices.

Por analogía con las formas bilineales, se llamará *rango* de una forma cuadrática el rango de su matriz en cualquier base. Si la matriz de una forma cuadrática es regular, la forma cuadrática se denominará también *regular*.

El estudio de las formas bilineales significa, en esencia, el estudio de sus matrices en diferentes bases o, lo que es igual, el estudio de la clase de matrices congruentes. Por ello, las investigaciones que siguen quedarán relacionadas con el examen de las clases de las matrices congruentes y congruentes según Hermite.

Para las clases de este género pueden indicarse ahora mismo toda una serie de propiedades que provienen de los resultados obtenidos anteriormente. Así por ejemplo, una matriz congruente de la simétrica (antisimétrica) es necesariamente simétrica (antisimétrica). En particular, una matriz congruente de la matriz diagonal será simétrica. De aquí concluimos que una matriz simétrica no nula nunca será congruente de la antisimétrica, aunque puede ser equivalente a ella. Una matriz antisimétrica no nula nunca puede ser congruente de la diagonal. Una matriz, congruente según Hermite de una matriz hermitiana (antihermitiana), es obligatoriamente hermitiana (antihermitiana). Entre las matrices diagonales, como matriz hermitiana (antihermitiana) puede intervenir sólo la que tiene elementos reales (imaginarios puros).

En concordancia con las descomposiciones (90.4), (90.7) de las formas bilineales y bilineales según Hermite, obtenemos unas descomposiciones de una matriz arbitraria en la suma de las matrices simétrica y antisimétrica, así como también de las matrices hermitiana y antihermitiana. Estas descomposiciones pueden ser escritas en la forma explícita:

$$A=\frac{1}{2}(A+A')+\frac{1}{2}(A-A'),$$

$$A=\frac{1}{2}(A+A^*)+\frac{1}{2}(A-A^*).$$

Si A es una matriz de la forma bilineal, los primeros sumandos de los segundos miembros son matrices de las partes simétricas de la forma bilineal y los segundos, las matrices de las partes antisimétricas de la misma forma.

En adelante haremos extender frecuentemente, sin explicaciones adicionales, la terminología introducida para las formas bilineales y cuadráticas a las matrices. Por ejemplo, llamemos una matriz *definida positiva*, entendiendo por ello que es una matriz de una forma definida positiva, etc.

Uno de los problemas más importantes relacionados con la forma bilineal es la determinación de la expresión más simple a la que puede ser reducida la matriz de la forma citada cuando varía la base y la busqueda de la base correspondiente. Este problema lleva el nombre de *transformación* de la forma bilineal o de *reducción* de la forma bilineal a una expresión más simple.

En la interpretación matricial el problema de transformación puede enunciarse de la manera siguiente:

Dada la matriz A, hállese tal matriz regular P que la matriz

$$C = P'AP, \tag{91.9}$$

congruente de A, tenga la forma más simple.

Esto nos da, de hecho, una descomposición de la matriz en factores, puesto que de (91.9) se deduce que

$$A = (P^{-1})'\, CP^{-1}.$$

Por supuesto, para las formas bilineales hermitianas, en lugar de (91.9) consideraremos las transformaciones

$$C=P'A\overline{P}. \tag{91.10}$$

Desde el punto de vista de los cálculos es importante que la matriz P en (91.9), (91.10) sea no muy compleja. Esto se debe a que al buscar las nuevas coordenadas de los vectores en términos de las antiguas, conforme a (63.3), nos vemos obligados a resolver un sistema de ecuaciones algebraicas lineales con la matriz P y es menester que la resolución se realice lo suficientemente rápido. En algunos casos, en lugar de la matriz P resulta más cómodo buscar la matriz P^{-1}.

Además de las formas consideradas de escribir las formas bilineales y cuadráticas, se usan también algunas otras. A veces las definiremos explícitamente:

$$\Phi = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} x_i y_j,$$

$$F = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} x_i x_j. \tag{91.11}$$

Estas anotaciones pueden ser simplificadas. Sea, por ejemplo, real un espacio, entonces reales serán también tanto la propia forma bilineal como la matriz A compuesta por los coeficientes a_{ji}. Introduzcamos el espacio $\mathbf{R}_n$ cuyos elementos son los vectores columna

$$x = (x_1, x_2, \ldots, x_n)', \quad y = (y_1, y_2, \ldots, y_n)'$$

y supongamos que el producto escalar se ha introducido como una suma de productos de las coordenadas tomados dos a dos. Ahora podemos escribir

$$\Phi = (Ax, y), \quad F = (Ax, x). \tag{91.12}$$

Para las formas bilineales hermitianas, anotadas como

$$\Phi = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i \overline{y}_j, \qquad F = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} x_i \overline{x}_j,$$

se verifica también (91.12), si, desde luego, el producto escalar se introduce como una suma de productos de las coordenadas del primer vector por las coordenadas conjugadas complejas del segundo vector.

Ejercicios.

1. Demuéstrese que el determinante de una matriz hermitiana es un número real.
2. ¿Qué número es el determinante de una matriz antihermitiana?
3. Demuéstrese que el rango de una matriz antisimétrica es un número par.
4. Las formas bilineales $\varphi(x, y)$ y $\overline{\varphi(y, x)}$ son, en general, diferentes. ¿Qué puede decirse sobre sus matrices?
5. Demuéstrese que el rango de la suma de unas formas bilineales no es superior a la suma de los rangos de los sumandos.
6. Demuéstrese que se puede representar toda forma bilineal de rango r como suma de r formas bilineales de rango 1.
7. Demuéstrese que se puede representar toda forma bilineal $\varphi(x, y)$ de rango 1 como

$$\varphi(x, y) = \varphi(x, a) \cdot \varphi(b, y)$$

para ciertos vectores a, b. ¿Será única tal representación?

§ 92. Reducción a una forma canónica

Antes de empezar a investigar las diferentes esferas del empleo de las formas bilineales y cuadráticas, consideraremos un método general de la transformación, congruente o congruente según Hermite, de las matrices a una forma sencilla.

Sea dada una matriz cuadrada A de orden n y se necesita hallar tal matriz regular P que la matriz $C = P'AP$ tenga una forma suficientemente sencilla. Realizándose una transformación congruente según Hermite, una forma sencilla la debe poseer la matriz $C = P'A\bar{P}$. Expondremos ahora un método general de la transformación que será útil para todas las matrices A. La diferencia entre las transformaciones congruente y congruente según Hermite será insignificante. Por ello, para concretar, convengamos en considerar que se realiza la transformación congruente de una matriz.

El método consiste en construcción de una sucesión de matrices $A_0 = A, A_1, A_2, \ldots, A_s$ en la que cada matriz consecutiva sea congruente de la anterior, es decir,

$$A_{k+1} = P'_{k+1} A_k P_{k+1}$$

para cierta matriz P_{k+1}. Puesto que la relación de congruencia es transitiva, la última matriz A_s será congruente de la matriz inicial A. El principio de construcción de la sucesión de matrices A_k está fundado en que para todo k se obtengan en la matriz A_{k+1} más elementos nulos que en la matriz A_k. Más aún, cada vez, al calcular la matriz P_{k+1} según la A_k, exigiremos que en la matriz A_{k+1} no sólo aparezcan nuevos elementos nulos, sino que se guarden todos los elementos nulos obtenidos en todas las etapas antecedentes.

La transformación de una matriz A_k en la A_{k+1} se llamará paso principal del método. Cada paso principal puede consistir en varios pasos auxiliares. Todos ellos se reducirán a la ejecución de las operaciones elementales: la permutación de las columnas (filas) de una matriz, la adición a una columna (fila) de otra columna (fila) multiplicada por un número, la multiplicación de una columna (fila) por un número. Describiremos los pasos auxiliares en términos de las transformaciones de la matriz A en otra matriz, congruente de ella, $C = P'AP$, omitiendo, para simplificar, el índice k.

A. En la matriz A el elemento $a_{11} \neq 0$. Existe una matriz regular P tal que para los elementos de la primera columna de la matriz $C = P'AP$ se verifican las correlaciones

$$c_{j1} = \begin{cases} a_{11}, & j = 1, \\ 0, & j \neq 1. \end{cases} \tag{92.1}$$

La matriz P difiere de la matriz unidad sólo en su primera fila, con la particularidad de que

$$p_{1j}=\begin{cases} 1, & j=1, \\ -\dfrac{a_{j1}}{a_{11}}, & j\neq 1. \end{cases} \tag{92.2}$$

La multiplicación a la izquierda de la matriz A por P' no altera la primera fila de la matriz A y convierte en cero todos los elementos de la primera columna de la matriz $P'A$ dispuestos fuera de la diagonal. La multiplicación a la derecha de la matriz $P'A$ por P no cambia la primera columna de la matriz $P'A$.

Hemos de señalar una circunstancia más. Llamaremos *principales* a todos los menores de la matriz dispuestos en la esquina izquierda superior. Como la matriz P es triangular derecha y todos los elementos diagonales de ella son iguales a la unidad, de todos los menores dispuestos en las primeras r columnas será distinto de cero sólo el menor principal: es igual a la unidad. Por esta razón en las matrices A y C coincidirán todos los menores principales. En efecto, haciendo uso de la fórmula Binet—Cauchy, obtenemos

$$C\begin{pmatrix}1 & 2\dots r\\ 1 & 2\dots r\end{pmatrix}=\sum_{1\leqslant k_1<k_2<\dots<k_r\leqslant n} P'\begin{pmatrix}1 & 2\dots r\\ k_1 & k_2\dots k_r\end{pmatrix}\times$$

$$\times AP\begin{pmatrix}k_1 & k_2\dots k_r\\ 1 & 2\dots r\end{pmatrix}=AP\begin{pmatrix}1 & 2\dots r\\ 1 & 2\dots r\end{pmatrix}=$$

$$=\sum_{1\leqslant k_1<k_2\dots<k_r\leqslant n} A\begin{pmatrix}1 & 2\dots r\\ k_1 & k_2\dots k_r\end{pmatrix}\cdot P\begin{pmatrix}k_1 & k_2\dots k_r\\ 1 & 2\dots r\end{pmatrix}=$$

$$=A\begin{pmatrix}1 & 2\dots r\\ 1 & 2\dots r\end{pmatrix}.$$

Esta observación la emplearemos más adelante.

B. En la matriz A el elemento a_{11} es igual a cero, pero cierto elemento a_{jj} es distinto de 0, $j>1$. Existe una matriz regular P tal que para la matriz $C=P'AP$ el elemento $c_{11}=a_{jj}$ es diferente de cero. La matriz P se diferencia de la matriz unidad sólo en cuatro elementos dispuestos en la intersección de las filas y las columnas con números 1, j. En estas posiciones la matriz P tiene por expresión $\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$. La multiplicación a la derecha de la matriz A por P permuta en la matriz A las columnas con los números 1, j. La multiplicación de la matriz AP a la izquierda por la matriz P' permuta en la matriz AP las filas con los números 1, j.

C. En la matriz A todos los elementos diagonales son nulos, pero hay tales índices j, l, donde $j<l$, que $a_{jl}+a_{lj}\neq 0$. Existe una

matriz regular P tal que para la matriz $C = P'AP$ el elemento $c_{jj} = a_{lj} + a_{jl}$ es distinto de 0. La matriz P se diferencia de la matriz unidad en el elemento $p_{lj} = 1$. La multiplicación a la derecha de la matriz A por P agrega a la j-ésima columna de la matriz A su l-ésima columna. La multiplicación a la izquierda de la matriz AP por P' agrega a la j-ésima fila de la matriz AP su l-ésima fila.

D. La matriz A es antisimétrica no nula, el elemento a_{12} es igual a 0, pero cierto elemento a_{jl} es distinto de 0, donde $j < l$. Existe una matriz regular P tal que en la matriz antisimétrica $C = P'AP$ el elemento $c_{12} = a_{jl}$ es diferente de 0. La matriz P viene representada como el producto $P = P_1 \cdot P_2$. Las matrices P_1, P_2 se diferencian de las matrices unidad sólo en cuatro elementos dispuestos en la intersección de las filas y las columnas con los números respectivos 1, j y 2, l. En estas posiciones las matrices P_1 y P_2 tienen por expresión $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Como ya se ha dicho, la multiplicación a la derecha por estas matrices conlleva la permutación de las columnas, la multiplicación a la izquierda, la permutación de las filas.

E. La matriz del menor principal de tercer orden de la matriz A tiene por expresión

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{pmatrix}, \tag{92.3}$$

donde los elementos a_{11}, a_{23} y a_{32} son distintos de cero. Existe una matriz regular P tal que en la matriz $C = P'AP$ serán diferentes de cero los primeros tres menores principales. La matriz P se diferencia de la matriz unidad en un elemento p_{31} el cual puede ser cualquier número, salvo 0, $-a_{12}a_{32}^{-1}$ y $-a_{11}a_{13}^{-1}$. La multiplicación a la derecha de la matriz A por P agrega a la primera columna de la matriz A su tercera columna multiplicada por p_{31}. La multiplicación a la izquierda de la matriz AP por P' agrega a la primera fila de la matriz AP su tercera fila multiplicada por p_{31}.

F. La matriz A es antisimétrica, el elemento a_{12} es distinto de 0. Existe una matriz regular P tal que para los elementos de las primeras dos columnas de la matriz $C = P'AP$ se verifican las correlaciones

$$c_{j1} = \begin{cases} -a_{12}, & j = 2, \\ 0 & j \neq 2, \end{cases} \quad c_{j2} = \begin{cases} a_{12}, & j = 1, \\ 0, & j \neq 1. \end{cases}$$

Puesto que en una transformación congruente una matriz antisimétrica se transforma en otra, también antisimétrica, las correlaciones análogas tendrán lugar también para las primeras dos filas de la matriz C. La matriz P se representa como un producto $P = P_1 \cdot P_2$. La matriz P_1 se diferencia de la matriz unidad sólo en la segunda fila, siendo, además,

$$p_{2j}^{(1)} = \begin{cases} 0, & j=1, \\ 1, & j=2, \\ \frac{a_{j1}}{a_{12}}, & j>2. \end{cases}$$

La matriz P_2 se diferencia de la matriz unidad sólo en la primera fila, verificándose en este caso.

$$p_{1j}^{(2)} = \begin{cases} 1, & j=1 \\ 0, & j=2, \\ -\frac{a_{j2}}{a_{12}}, & j>2. \end{cases}$$

La multiplicación a la izquierda de la matriz A por P_1' no altera las primeras dos filas y la segunda columna de la matriz A, convirtiendo en cero todos los elementos en la primera columna de la matriz $P_1'A$, a excepción de los dos primeros. La multiplicación a la izquierda de la matriz $P_1'A$ por P_2' no altera las primeras dos filas y la primera columna de la matriz $P_1'A$, convirtiendo en cero todos los elementos de la segunda columna de la matriz $P'A$, a excepción de los dos primeros. La multiplicación a la derecha de la matriz $P'A$ por P no cambia las primeras dos columnas de la matriz $P'A$.

G. Supongamos que la matriz A tiene, realizada cierta partición en células, la siguiente estructura

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array} \right), \tag{92.4}$$

donde A_{11}, A_{12} son células cuadradas. Si P_{22} es una matriz regular cuyo orden es igual al de A_{22}, entonces la matriz

$$C = \left(\begin{array}{c|c} A_{11} & A_{12}P_{22} \\ \hline 0 & P_{22}'A_{22}P_{22} \end{array} \right)$$

es congruente de la matriz A. En este caso $C = P'AP$, donde

$$P = \left(\begin{array}{c|c} E & 0 \\ \hline 0 & P_{22} \end{array} \right).$$

La comprobación directa de todas las afirmaciones enunciadas al discribir los pasos auxiliares no representa alguna dificultad singular y por esto proponemos que el lector mismo se convenza de su veracidad lo que puede hacerse en calidad de ejercicios.

El método, en total, se realiza del modo siguiente. En el primer paso principal la matriz A se reduce a la forma (92.4), donde A_{11} es una matriz regular de orden uno o dos. Si la matriz A_k, $k \geqslant 1$, tiene la forma (92.4), entonces, al realizarse el segundo paso principal, la matriz en la esquina inferior derecha se reduce también a la forma (92.4) y se lleva a cabo la transformación congruente general de conformidad con el paso **G**. La matriz A_{k+1} puede, nuevamente, representarse en la forma (92.4), pero la célula en la esquina izquierda superior para A_{k+1} no sólo será regular, sino tendrá el orden mayor en comparación con el de la matriz A_k. El proceso se repite hasta que, realizado algún paso, en la representación (92.4) de la matriz A_s aparezca una célula nula en la esquina inferior derecha o bien el orden de la célula en la esquina superior izquierda se haga igual a n. En este caso, la matriz de la transformación resultante será igual al producto de izquierda a derecha de las matrices de transformaciones de todos los pasos.

La forma de la matriz A_s depende de si es o no la matriz A antisimétrica. De esto mismo depende también de qué modo los pasos auxiliares forman parte de los pasos principales del método.

Cualquiera que sea el paso principal, su finalidad consiste en obtener una porción seguida de ceros en la matriz a transformar. Si la matriz inicial no es antisimétrica, los ceros se obtienen siempre con ayuda del paso auxiliar **A**, mientras que los pasos **B** — **C** sólo se necesitan para preparar **A**. En cambio, si la matriz inicial es antisimétrica, los ceros se obtienen con ayuda del paso **F**, mientras que **D** es el paso auxiliar. Describamos también el paso principal del método en términos de la transformación de la matriz A y empecemos con la matriz no antisimétrica A.

En el primer paso principal, la matriz que se transforma no es antisimétrica. Si el elemento $a_{11} \neq 0$ y todos los elementos extradiagonales de la primera columna son nulos, entonces nada varía y consideramos terminado el paso principal. A título de matriz de la transformación P tomamos, en este caso, la matriz unidad. En el caso general realizamos el primero de los pasos auxiliares **A** — **C**, que puede ser llevado a cabo. Si tal paso resulta ser **B** o **C**, después de éste se cumple obligatoriamente el paso **A** o bien ambos pasos **B**, **A**. A título de matriz de la transformación P tomamos el producto de izquierda a derecha de todas las matrices de transformaciones de los pasos auxiliares realmente realizados. Después de cumplir el primer paso principal, en la matriz transformada A_1 todos los elementos extradiagonales de la primera columna serán nulos, es decir, la matriz A_1 será de estructura celular del tipo (92.4).

La diferencia de todos los pasos restantes con respecto al primero está relacionada con el hecho de que la matriz a transformar puede resultar antisimétrica. Si ésta no es antisimétrica, el paso principal siguiente no difiere en nada del paso primero. En cambio, si la matriz que se transforma es antisimétrica, entonces cualquiera que sea su transformación congruente, queda antisimétrica y, sirviéndose sólo de esta matriz, no se puede obtener un elemento no nulo en la esquina superior izquierda. La salida de esta situación está basada en la necesidad de transformar la célula diagonal inferior ampliada.

Hasta que se encuentre una matriz antisimétrica, la célula en la esquina superior izquierda de la representación (92.4) para las matrices A_k será triangular derecha con elementos diagonales no nulos. Si los elementos en las posiciones (1, 2) y (2, 1) de la matriz antisimétrica en la esquina inferior derecha son distintos de cero, entonces para la matriz A_k, la siguiente en la transformación, sustituyamos la representación (92.4), disminuyendo en uno el orden de la célula en la esquina superior izquierda. Ahora la matriz de tercer orden en la esquina superior izquierda de la nueva célula diagonal inferior tendrá la forma (92.3) y se puede realizar el paso auxiliar **E**. Después de esto podemos realizar tres veces consecutivas el paso **A**. Efectivamente, según lo observado, la ejecución del paso **A** no cambia los menores principales de la matriz. Por consiguiente, en el caso dado, realizado el paso **A**, la nueva matriz en la esquina inferior derecha tendrá diferentes de cero los dos primeros menores principales. Por ello, podemos, a ciencia cierta, hacer un paso **A** más. Los razonamientos análogos demuestran que el paso **A** puede realizarse también por tercera vez. Al retroceder un paso "atrás" hemos obtenido la posibilidad de avanzar tres pasos "adelante". Cuando sea necesario, antes de realizar el paso **E**, se lleva a cabo el paso **D**.

De suerte, si la matriz A no es antisimétrica, el método que acabamos de exponer permite construir una matriz regular P tal que la matriz $P'AP$, congruente de A, tendrá la estructura siguiente:

$$P'AP=\left(\begin{array}{c|c} M & N \\ \hline 0 & 0 \end{array}\right). \tag{92.5}$$

Aquí M es una matriz triangular derecha con los elementos diagonales no nulos, el orden de la matriz M es igual al rango de la matriz A.

Si A es una matriz antisimétrica, todos los pasos principales del método, incluido el primero, se realizan siguiendo un mismo esquema. Supongamos que ya se ha obtenido la matriz A_k del tipo (92.4), con la particularidad de que en la esquina superior izquierda se dispone una matriz diagonal celular regular de células antisimétricas de segundo orden. Dado que, al realizar una transformación congruente, una matriz antisimétrica se transforma en otra antisimétrica, entonces la célula A_{12} en (92.4) será nula. Primero se logra que en las posiciones (1, 2) y (2, 1) de la matriz antisimétrica la esquina inferior derecha sea ocupada por elementos no nulos. Es posible que para esto resulte necesario realizar el paso auxiliar **D**. Luego realizamos el paso **F**, lo que adjunta a la diagonal una célula más que es antisimétrica regular de segundo orden. A continuación, pasamos al siguiente paso principal. En este caso también el proceso continúa hasta que, realizado cierto paso, en la representación (92.4) de la matriz A_s aparezca una célula nula en la esquina inferior derecha, o bien el orden de la célula en la esquina superior izquierda se haga igual a n.

Así pues, si la matriz A es antisimétrica, el método permite construir una matriz regular P, para la cual la matriz $P'AP$ será de la estructura siguiente:

$$P'AP=\left(\begin{array}{c|c} M & 0 \\ \hline 0 & 0 \end{array}\right) \tag{92.6}$$

Aquí M es una matriz diagonal celular con células antisimétricas regulares de segundo orden. El orden de la matriz M es igual al rango de la matriz A.

En la transformación congruente hermitiana el esquema general del método queda inalterable. Sin embargo, el propio proceso resulta ser más fácil, comparado con la transformación congruente ordinaria, si el paso auxiliar C se sustituye por el siguiente.

C'. En la matriz A todos los elementos diagonales son nulos, pero hay tales índices j, l, donde $j < l$, que entre los elementos a_{lj}, a_{jl} existe aunque sea uno diferente de cero. Existe una matriz regular P tal que para la matriz $C = P'AP$ uno de los elementos diagonales c_{jj}, c_{ll} es distinto de cero. A saber, $c_{jj} = a_{jl} + a_{lj}$, $c_{ll} = i\,(a_{jl} - a_{lj})$. La matriz P se diferencia de la matriz unidad en dos elementos $p_{lj} = 1$, $p_{jl} = i$. La multiplicación a la derecha de la matriz A por P agrega a la j-ésima columna de la matriz A su l-ésima columna y a la l-ésima columna, su j-ésima columna multiplicada por $-i$. La multiplicación a la izquierda de la matriz $A\overline{P}$ por P' adjunta a la j-ésima fila de la matriz $A\overline{P}$ su l-ésima fila y a la l-ésima fila, su j-ésima fila multiplicada por i.

Ahora no hay necesidad en los pasos **D** — **F** del método general, puesto que nunca sobrepasaremos el paso **C'**. Además, las fórmulas (92.2) quedan intactas.

De este modo, si A es una matriz no nula, el método hace posible construir tal matriz regular P que la matriz $P'A\overline{P}$, congruente de A según Hermite, será de la siguiente estructura:

$$P'A\overline{P}=\left(\begin{array}{c|c} M & N \\ \hline 0 & 0 \end{array}\right). \tag{92.7}$$

Aquí, M es una matriz triangular derecha con elementos diagonales no nulos. El orden de la matriz M es igual al rango de la matriz A.

Los tipos de las matrices (92.5) — (92.7) se denominan formas *canónicas* para las operaciones de la transformación congruente. Se llama *canónica* también cualquier base en la que la matriz inicial tiene la forma indicada. Las propias matrices del tipo (92.5), (92.7) llevan el nombre de *trapezoidales derechas*. De modo análogo se definen las matrices *trapezoidales izquierdas*.

Demos a conocer algunas deducciones interesantes que provienen de las formas canónicas de matrices. Ya hemos dicho que en una transformación congruente se conserva el carácter simétrico y antisimétrico de la matriz. Si una de estas propiedades la poseía la matriz inicial, debe quedarse válida para la forma canónica. Por esto, en adición a lo dicho podemos concluir que

Una matriz simétrica es congruente de la matriz diagonal.

Una matriz hermitiana es congruente según Hermite de la matriz diagonal real.

Una matriz antihermitiana es congruente según Hermite de la matriz diagonal imaginaria pura.

En todos estos casos la reducción a una forma canónica se efectúa con una facilidad singular, puesto que no puede surgir la necesidad de llevar a cabo aunque sea uno solo de los pasos auxiliares **D** — **F**.

Sobre las matrices de la forma canónica del tipo (92.5), (92.6) se puede realizar una transformación congruente con la matriz diagonal más y conseguir que los elementos no nulos que determinan la regularidad de la célula M sean iguales a $+1$ o bien a -1. Esta forma canónica de la matriz y la base que le corresponde se denominan *normales*. Está claro que la multiplicación a la derecha (a la izquierda) por una matriz diagonal conduce a la multiplicación de las columnas (filas) por los elementos diagonales de la matriz de la transformación. Describamos nuevamente esta transformación en términos del cumplimiento del paso auxiliar con la matriz A.

H. Una matriz real no antisimétrica A de rango r es de la forma canónica (92.5). Existe una matriz diagonal real P tal que los elementos diagonales no nulos c_{jj} de la matriz $C = P'AP$ son iguales a sign a_{jj}. Además

$$p_{jj} = \begin{cases} (a_{jj}\operatorname{sign} a_{jj})^{-1/2}, & j \leqslant r, \\ 1, & j > r, \end{cases}$$

Una matriz real (compleja) antisimétrica A de rango r es de la forma canónica (92.6). Existe una matriz diagonal real (compleja) P tal que los elementos no nulos dispuestos por arriba de la diagonal de la matriz $C = P'AP$ son iguales a $+1$, y los elementos no nulos dispuestos por debajo de la diagonal, iguales a -1. En este caso

$$p_{jj} = \begin{cases} 1, & j \text{ es impar}, \\ a_{j-1,\,j}^{-1}, & j \text{ es par}. \end{cases}$$

Una matriz compleja no antisimétrica A de rango r tiene la forma canónica (92.5). Existe una matriz diagonal compleja P tal que los elementos diagonales no nulos c_{jj} de la matriz $C = P'AP$ son iguales a uno. En este caso

$$p_{jj}=\begin{cases} a_{jj}^{-1/2}, & j\leqslant r, \\ 1, & j>r. \end{cases}$$

La transformación congruente según Hermite con una matriz diagonal se efectúa raras veces, puesto que con su ayuda sólo pueden cambiarse los módulos de los elementos que determinan la regularidad de la célula M en (92.7), pero no se puede hacer reales los elementos diagonales complejos.

Ejercicios.

1. Demuéstrese que si la reducción a la forma canónica mediante la matriz P se efectúa según el método descrito más arriba, entonces $\det P = \pm 1$.

2. ¿Qué significa, desde el punto de vista de la forma canónica, la igualdad matricial

$$\begin{pmatrix} 1 & 1 \\ 1 & -i \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -i \end{pmatrix}=\begin{pmatrix} 1+i & 0 \\ 2 & 1+i \end{pmatrix}? \tag{92.8}$$

3. ¿A qué forma puede reducirse una matriz no antisimétrica con ayuda de una transformación congruente, si se excluye el paso auxiliar E?

4. ¿Qué forma tiene la matriz P de una transformación, si cada paso principal del método considerado más arriba consistía sólo en el paso auxiliar A?

5. Demuéstrese que toda matriz triangular derecha es congruente de la matriz triangular izquierda. ¿Cuál es la forma más simple de la matriz de una transformación?

6. Demuéstrese que toda matriz regular de orden impar es congruente de la matriz triangular derecha regular.

7. Sea G una matriz de una forma bilineal definida positiva. Demuéstrese que para sus elementos g_{ij} se verifican las correlaciones

$$g_{ii}>0, \qquad (g_{ij}+g_{ji})^2<4g_{ii}g_{jj},$$

cualesquiera que sean i, j.

8. Sea G una matriz de una forma bilineal definida negativa. Demuéstrese que para sus elementos g_{ij} se verifican las correlaciones

$$g_{ii}<0, \qquad (g_{ij}+g_{ji})^2<4g_{ii}g_{jj},$$

cualesquiera que sean i, j.

9. Demuéstrese que las matrices de todas las formas bilineales simétricas definidas positivas (negativas) son congruentes entre sí mismas.

10. Demuéstrese que para que G sea una matriz de la forma bilineal de signo variable, es suficiente que entre sus elementos diagonales haya elementos de signos distintos.

§ 93. Congruencia y descomposiciones matriciales

El método general de la transformación congruente de una matriz a la forma canónica no siempre permite decir de antemano, cuál será la matriz de la transformación de coordenadas al pasar a la base canónica. No obstante, con ciertas restricciones adicionales impuestas en la matriz inicial, para dicha pregunta existe una respuesta bien determinada.

Supongamos que la matriz A tiene distintos de cero todos los menores principales, a excepción, quizás, del menor de orden superior, es decir, del determinante de la matriz A. Probemos que tal matriz siempre puede representarse en forma del producto

$$A = LDU, \tag{93.1}$$

donde L es una matriz triangular izquierda con elementos diagonales unidades, D es una matriz diagonal y U, una matriz triangular derecha con elementos diagonales unidades, es decir,

$$A = \begin{pmatrix} 1 & & \\ l_{21} & 1 & 0 \\ l_{n1} & l_{n2} \dots & 1 \end{pmatrix} \begin{pmatrix} d_{11} & & \\ & d_{22} & 0 \\ 0 & & d_{nn} \end{pmatrix} \begin{pmatrix} 1 & u_{12} \dots & u_{1n} \\ & 1 \dots & u_{2n} \\ 0 & & 1 \end{pmatrix}.$$

Igualando entre sí los elementos de la matriz A y del producto LDU obtenemos

$$a_{ij} = \sum_{p=1}^{\min(i,\,j)} l_{ip} d_{pp} u_{pj}. \tag{93.2}$$

Ahora, de (93.2) hallamos de manera sucesiva todos los elementos desconocidos de las matrices de la descomposición (93.1). A saber,

$$\begin{gathered} d_{11} = a_{11}, \\ u_{1j} = \frac{a_{1j}}{d_{11}}, \quad l_{j1} = \frac{a_{j1}}{d_{11}}, \quad j > 1, \\ d_{ii} = a_{ii} - \sum_{p=1}^{i-1} l_{ip} d_{pp} u_{pi}, \quad i > 1, \\ u_{ij} = \frac{a_{ij} - \sum\limits_{p=1}^{i-1} l_{ip} d_{pp} u_{pj}}{d_{ii}}, \\ l_{ji} = \frac{a_{ji} - \sum\limits_{p=1}^{j-1} l_{jp} d_{pp} u_{pi}}{d_{ii}}, \quad i > 1, \quad j > i. \end{gathered} \tag{93.3}$$

Apliquemos a la correlación (93.1) la fórmula de Binet—Cauchy. Recordemos que entre los menores de la matriz triangular izquierda L, dispuestos en las primeras r filas, sólo el menor principal es diferente de cero: es igual a la unidad. Los razonamientos análogos tienen lugar también para la matriz U, al cambiar entre sí, por supuesto, las filas y las columnas. Por ello

$$A\begin{pmatrix}1 & 2 \dots r \\ 1 & 2 \dots r\end{pmatrix} = \sum_{1 \leqslant k_1 < k_2 \dots < k_r \leqslant n} L\begin{pmatrix}1 & 2 & \dots & r \\ k_1 & k_2 & \dots & k_r\end{pmatrix} \times$$

$$\times DU\begin{pmatrix}k_1 & k_2 \dots & k_r \\ 1 & 2 \dots & r\end{pmatrix} = DU\begin{pmatrix}1 & 2 \dots r \\ 1 & 2 \dots r\end{pmatrix} = d_{11}d_{22} \dots d_{rr}.$$

De aquí concluimos que

$$d_{11} = a_{11}, \qquad d_{ii} = \frac{A\begin{pmatrix}1 & 2 \dots i \\ 1 & 2 \dots i\end{pmatrix}}{A\begin{pmatrix}1 & 2 \dots i-1 \\ 1 & 2 \dots i-1\end{pmatrix}}. \tag{93.4}$$

Por hipótesis, los menores principales de la matriz A son distintos de cero. Por consiguiente, serán distintos de cero todos los elementos diagonales d_{ii} en (93.4), salvo, quizás, el último elemento.

Trataremos con frecuencia las descomposiciones (93.1) para las matrices simétricas y hermitianas. Si esta vez también la matriz A tiene sus menores principales distintos de cero, a excepción, quizás, del último, entonces la matriz simétrica siempre puede ser representada en forma del producto

$$A = S'DS, \tag{93.5}$$

y la matriz hermitiana, en forma del producto

$$A = S'D\bar{S}. \tag{93.6}$$

Aquí S es una matriz triangular derecha cuyos elementos diagonales son iguales a uno, D es una matriz diagonal, es decir,

$$S = \begin{pmatrix}1 & s_{12} \dots s_{1n} \\ & 1 \dots s_{2n} \\ & \ddots \\ 0 & \quad 1\end{pmatrix} \qquad D = \begin{pmatrix}d_{11} & & 0 \\ & d_{22} & \\ & & \ddots \\ 0 & & d_{nn}\end{pmatrix}$$

En concordancia completa con (93.3) tendremos ahora

$$d_{11} = a_{11}, \quad s_{1j} = \frac{a_{1j}}{d_{11}}, \quad j > 1,$$

$$d_{ii} = a_{ii} - \sum_{p=1}^{i-1} d_{pp}s_{pi}^2, \quad i > 1, \tag{93.7}$$

$$s_{ij} = \frac{a_{ij} - \sum_{p=1}^{i-1} d_{pp}s_{pi}s_{pj}}{d_{ii}}, \quad j > i,$$

para la descomposición (93.5) y

$$d_{11}=a_{11}, \quad s_{1j}=\frac{\bar{a}_{1j}}{d_{11}}, \quad j>1,$$

$$d_{ii}=a_{ii}-\sum_{p=1}^{i-1} d_{pp}\,|s_{pi}|^2, \quad i>1,$$

$$s_{ij}=\frac{a_{ij}-\sum\limits_{p=1}^{i-1} d_{pp}s_{pi}\bar{s}_{pj}}{d_{ii}}, \quad j>i,$$

para la descomposición (93.6). En este caso las fórmulas (93.4) siguen siendo válidas.

Las descomposiciones (93.1), (93.5), (93.6) son de amplio uso en la resolución de los más diversos problemas del álgebra lineal. En lo que se refiere a las transformaciones congruentes de una matriz, los datos de la descomposición conducen a las siguientes correlaciones:

$$(L^{-1'})'\,AL^{-1'}=DUL^{-1'}, \qquad (L^{-1'})'\,A\overline{L^{-1'}}=DU\overline{L^{-1'}},$$

$$S^{-1'}\,AS^{-1}=D, \qquad S^{-1'}A\bar{S}^{-1}=D.$$

Las matrices $DUL^{-1'}$ y $DU\overline{L^{-1'}}$ son triangulares derechas, las D son diagonales, con la particularidad de que en sus diagonales principales solamente el último elemento puede ser nulo. De suerte que otra vez hemos obtenido los tipos ya conocidos de las matrices en una transformación congruente. Sin embargo, ahora se puede afirmar que las matrices de la transformación de coordenadas, al pasar a la base canónica, serán triangulares derechas, puesto que lo son las matrices $L^{-1'}$, S^{-1}. Las descomposiciones examinadas proporcionan las matrices L', S de las transformaciones de coordenadas, al pasar de la base canónica a la inicial, las cuales asimismo serán triangulares derechas.

En el caso de una matriz simétrica el proceso descrito de descomposición está estrechamente relacionado con el así llamado *algoritmo de Jacobi* de la transformación de una forma cuadrática en la forma canónica. La diferencia sólo consiste en que el algoritmo de Jacobi tiene determinada la matriz S^{-1}, en lugar de la matriz S. Ha de señalarse que la matriz S se halla de un modo más simple que S^{-1}.

Las transformaciones congruentes con una matriz triangular derecha son las más sencillas, no obstante lo suficientemente generales todavía para que puedan ser aplicadas a una clase amplia de matrices. Por esta razón causa un interés determinado la descripción de aquella clase de matrices que pueden reducirse a la forma canónica con ayuda de una transformación con la matriz triangular derecha.

LEMA 93.1. *Si una matriz rectangular A está representada en la forma celular*

$$A=\left(\begin{array}{c|c} B & Q \\ \hline R & T \end{array}\right), \tag{93.8}$$

donde B es una matriz cuadrada regular de orden r, entonces el rango de la matriz A es igual a r, si, y sólo si,

$$T=RB^{-1}Q. \tag{93.9}$$

DEMOSTRACION. Multipliquemos la matriz A a la izquierda por una matriz celular regular

$$V=\left(\begin{array}{c|c} E & 0 \\ \hline -RB^{-1} & E \end{array}\right),$$

donde las células correspondientes tienen las mismas dimensiones que en (93.8). Entonces

$$VA=\left(\begin{array}{c|c} B & Q \\ \hline 0 & T-RB^{-1}Q \end{array}\right).$$

Las matrices A y VA son de un mismo rango el cual será igual a r cuando, y sólo cuando, $T-RB^{-1}Q=0$.

Ahora podemos describir la clase buscada de matrices. Resulta estrechamente relacionada con las matrices del tipo (93.8), (93.9).

TEOREMA 93.1. *Para que una matriz no antisimétrica pueda ser reducida a la forma canónica mediante la transformación congruente con una matriz triangular derecha, es necesario y suficiente que el número de los primeros menores principales no nulos de la matriz A sea igual a su rango.*

DEMOSTRACIÓN. NECESIDAD. Supongamos que una matriz no antisimétrica A se reduce, con ayuda de la matriz triangular derecha P, a la forma canónica (92.5). Es evidente que el número de los primeros menores principales no nulos en la matriz A no puede ser superior al orden de la célula M. Aplicando la fórmula de Binet—Cauchy y teniendo presente que en las primeras columnas de la matriz P el menor no nulo está ausente, salvo el principal, obtenemos

$$A\begin{pmatrix}1 & 2 \dots s \\ 1 & 2 \dots s\end{pmatrix}\left\{P\begin{pmatrix}1 & 2 \dots s \\ 1 & 2 \dots s\end{pmatrix}\right\}^2=M\begin{pmatrix}1 & 2 \dots s \\ 1 & 2 \dots s\end{pmatrix}$$

para todo s no superior al orden de la matriz M. Como los menores principales de la matriz M y P son distintos de cero, el número de los primeros menores principales no nulos de la matriz A es igual a su rango.

SUFICIENCIA. Supongamos que el número de los primeros menores principales no nulos de la matriz A y el rango de ésta son iguales a r. Representemos la matriz A en la forma celular (93.8), donde el orden de la célula B es igual a r. Puesto que todos los menores principales de la matriz B son distintos de cero, entonces, de acuerdo con lo dicho más arriba, se la puede representar en la forma $B = LDU$, análogamente a (93.1). Construyamos una matriz celular

$$P = \left(\begin{array}{c|c} L^{-1'} & -B^{-1'}R' \\ \hline 0 & E \end{array}\right).$$

La comprobación directa muestra que

$$P'AP = \left(\begin{array}{c|c} DUL^{-1'} & L^{-1}(-BB^{-1'}R' + Q) \\ \hline 0 & 0 \end{array}\right).$$

La matriz $DUL^{-1'}$ es triangular derecha regular, la matriz P es triangular derecha regular y, por lo tanto, la matriz A se reduce a la forma canónica de un modo adecuado.

Para la transformación congruente de una matriz antisimétrica y la transformación congruente hermitiana de una matriz arbitraria las afirmaciones correspondientes se demuestran análogamente y aquí nos limitamos sólo a enunciarlas.

TEOREMA 93.2. *Para que una matriz antisimétrica A de rango r pueda ser reducida a la forma canónica mediante la transformación congruente con una matriz triangular derecha, es necesario y suficiente que el número de los primeros menores principales no nulos del orden par de la matriz A sea igual a $r/2$.*

TEOREMA 93.3. *Para que una matriz A pueda ser reducida a la forma canónica mediante la transformación congruente hermitiana con una matriz triangular derecha, es necesario y suficiente que el número de los primeros menores principales no nulos de la matriz A sea igual al rango de ésta.*

Las transformaciones de una matriz, tanto congruentes como congruentes según Hermite, no son, en el caso general, transformaciones de semejanza. No obstante, si para cierta clase de matrices P se verifica uno de los grupos de las correlaciones

$$PP' = P'P = E, \qquad PP^* = P^*P = E, \tag{93.10}$$

entonces, en este caso la transformación de congruencia se convierte en la de semejanza y para realizar las investigaciones se pueden utilizar los resultados obtenidos anteriormente referentes a la semejanza de matrices. Como ya sabemos, las matrices ortogonales reales satisfacen al primer grupo de las correlaciones en (93.10), las matrices unitarias complejas, al segundo grupo de correlaciones. Por eso, al recordar los resultados de los §§ 76—81, referentes a las semejanzas unitaria y ortogonal, concluimos que son lícitas las afirmaciones siguientes.

Toda matriz real, simétrica o antisimétrica, se reduce a la forma canónica mediante la transformación congruente con una matriz ortogonal.

Toda matriz compleja se reduce a la forma canónica por medio de la transformación congruente hermitiana con una matriz unitaria.

Estas afirmaciones son, en lo principal, de interés teórico, dado que en la práctica resulta muy difícil hallar matrices unitarias y ortogonales de la transformación, sobre todo cuando $n \geqslant 5$.

Ejercicios.

1. Demuéstrese que si las descomposiciones (93.1), (93.5), (93.6) existen, son únicas.

2. Demuéstrese que si todos los menores (a excepción, quizás, del menor de orden superior) de la matriz A, dispuestos en la esquina inferior derecha, son distintos de cero, entonces existe, y, además, sólo una descomposición $A = LDU$, donde L es una matriz triangular derecha, U, una matriz triangular izquierda con elementos diagonales iguales a uno, D, una matriz diagonal.

3. Demuéstrese que para los elementos d_{ii} de la matriz D que figura en el ejercicio 2, son válidas las correlaciones

$$d_{nn} = a_{nn}; \qquad d_{ii} = \frac{A\begin{pmatrix} i, & i+1, & \ldots, & n \\ i, & i+1, & \ldots, & n \end{pmatrix}}{A\begin{pmatrix} i+1, & i+2, & \ldots, & n \\ i+1, & i+2, & \ldots, & n \end{pmatrix}}, \quad i < n.$$

4. ¿En qué factores triangulares puede descomponerse una matriz, si son distintos de cero sus menores dispuestos en la esquina izquierda inferior (derecha superior)?

5. Supongamos que para los elementos a_{ij} de la matriz A se verifican las correlaciones

$$a_{ij} = 0, \qquad k < j - i, \; j - i < l, \tag{93.11}$$

para ciertos números $l < k$. Tal matriz se llama *matriz de cinta*. Demuéstrese que si para una matriz de cinta A tiene lugar la descomposición (93.1), entonces

$$l_{ij} = 0, \qquad j - i < l, \qquad u_{ij} = 0, \qquad j - i > k.$$

6. Una matriz A se llama *tridiagonal*, si satisface las condiciones (93.11) cuando $k = 1$, $l = -1$. ¿Qué forma tienen las fórmulas (93.3), (93.7) para la matriz tridiagonal?

7. Una matriz A se llama *casi triangular derecha* (*izquierda*), si satisface las condiciones (93.11) para $k = n$, $l = -1$ ($k = 1$, $l = -n$). ¿Qué forma tienen las fórmulas (93.3) para las matrices casi triangulares?

8. ¿Qué cantidad de operaciones aritméticas deben realizarse para diferentes tipos de matrices cuando se obtienen las descomposiciones del tipo (93.1)?

9. ¿Como deben aplicarse las descomposiciones (93.1), (93.5), (93.6) para la resolución de los sistemas de ecuaciones algebraicas lineales?

§ 94. Formas bilineales simétricas

Al examinar las formas bilineales y cuadráticas prestábamos más de una vez especial atención tanto a las formas bilineales simétricas como a las bilineales que engendran formas cuadráticas reales. Sólo dos tipos de las formas bilineales satisfacen simultáneamente ambas condiciones: la forma bilineal simétrica real y la forma bilineal simétrica hermitiana. Las matrices de estas formas son en cualquier base simétrica real o hermitiana, respectivamente. Ambos tipos de las matrices se reducen, por medio de una transformación congruente, a la forma normal real diagonal.

Según hemos visto, una misma matriz puede reducirse a la forma canónica mediante diferentes transformaciones congruentes. Por eso, en el caso general, la forma canónica no es unívocamente definida. Surge naturalmente una cuestión ¿qué tienen en común las diferentes formas canónicas a las cuales se reduce una misma matriz? Se sabe que el rango de una matriz no depende de la transformación, razón por la cual, cualquiera que sea el procedimiento de reducción a la forma canónica, el número de las últimas filas nulas será el mismo. Para las matrices simétrica real y hermitiana podemos decir mucho más. La forma canónica de estas matrices puede ser descrita por el número de sus términos positivos y negativos. Tiene lugar el siguiente teorema importante

TEOREMA 94.1 (*ley de inercia de las formas cuadráticas*). *El número de los términos positivos y el de los términos negativos en la forma canónica de una matriz simétrica real, en la transformación congruente ordinaria, y de una matriz hermitiana, en la transformación congruente hermitiana, no dependen de cómo se realiza la reducción.*

DEMOSTRACIÓN. Supongamos que una matriz A satisface las condiciones del teorema. Consideremos la forma cuadrática F con la matriz A de rango r de las variables $x_1, x_2, \ldots, x_n$ y supongamos que dicha matriz se ha reducido a la forma normal mediante dos procedimientos

$$F = y_1^2 + y_2^2 + \ldots + y_k^2 - y_{k+1}^2 - y_{k+2}^2 - \ldots - y_r^2 =$$
$$= z_1^2 + z_2^2 + \ldots + z_l^2 - z_{l+1}^2 - z_{l+2}^2 - \ldots - z_r^2. \qquad (94.1)$$

Puesto que el paso de las variables $x_1, x_2, \ldots, x_n$ a las variables $y_1, y_2, \ldots, y_n$ se ha realizado mediante una transformación lineal regular, las segundas variables se expresarán linealmente en términos de las primeras, con la particularidad de que el determinante de la matriz de la transformación inversa será diferente de cero. Así pues,

$$y_i = \sum_{s=1}^{n} b_{is} x_s, \qquad \det\begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \cdot & \cdot & \cdot \\ b_{n1} & \ldots & b_{nn} \end{pmatrix} \neq 0. \qquad (94.2)$$

Análogamente

$$z_j = \sum_{t=1}^{n} c_{jt} x_t, \qquad \det\begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \cdot & \cdot & \cdot \\ c_{n1} & \ldots & c_{nn} \end{pmatrix} \neq 0. \qquad (94.3)$$

Supongamos $k < l$ y escribamos un sistema de ecuaciones

$$y_1 = y_2 = \ldots = y_k = z_{l+1} = \ldots = z_n = 0. \qquad (94.4)$$

Si los primeros miembros de estas igualdades se sustituyen por sus expresiones de (94.2), (94.3), se obtiene un sistema de $n - l + k$ ecuaciones lineales homogéneas con n incógnitas $x_1, x_2, \ldots, x_n$. El número de ecuaciones en este sistema es menor que el número de incógnitas, a consecuencia de lo cual la solución del sistema es real no nula $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Sustituyamos ahora, en la igualdad (94.1), todas las variables por sus expresiones de (94.2), (94.3) y a continuación en lugar de las variables $x_1, x_2, \ldots, x_n$ escribamos los números $\alpha_1, \alpha_2, \ldots, \alpha_n$. Si, realizada tal sustitución, designamos, para brevedad, con $y_i(\alpha)$ y $z_j(\alpha)$ los valores de las variables y_i, z_j, entonces, teniendo en cuenta (94.4), la correlación (94.1) se convierte en la igualdad

$$-y_{k+1}^2(\alpha) - \ldots - y_r^2(\alpha) = z_1^2(\alpha) + \ldots + z_l^2(\alpha).$$

De aquí se desprende

$$z_1(\alpha) = \ldots = z_l(\alpha) = 0. \qquad (94.5)$$

Por otra parte, por la propia elección de los números $\alpha_1, \alpha_2, \ldots, \alpha_n$ tenemos

$$z_{l+1}(\alpha) = \ldots = z_r(\alpha) = \ldots = z_n(\alpha) = 0. \qquad (94.6)$$

De este modo, el sistema de n ecuaciones lineales homogéneas

$$z_i = 0, \qquad i = 1, 2, \ldots, n,$$

con n incógnitas $x_1, x_2, \ldots, x_n$ tiene, en virtud de (94.5), (94.6), una solución no nula $\alpha_1, \alpha_2, \ldots, \alpha_n$, es decir, el determinante de este sistema debe ser igual a cero. Esto contradice a (94.3). A una contradicción análoga llegaremos al suponer $l < k$. Por consiguiente, $l = k$ y el teorema queda demostrado.

Toda forma cuadrática real ordinaria (hermitiana) en un espacio lineal real (complejo) tiene en cualquier base una única matriz simétrica real (hermitiana compleja). Estas matrices satisfacen las condiciones del teorema 94.1. Cualquiera que sea la base el número de términos positivos y negativos de la forma canónica de una matriz es un invariante para la forma cuadrática y se denomina, respectivamente, su *índice de inercia* positivo y negativo. La diferencia entre el índice positivo y el negativo se denomina *signatura* de la forma cuadrática. Ahora se pueden enunciar algunos corolarios útiles del teorema 94.1.

COROLARIO. *Una forma cuadrática es definida positiva (negativa) cuando, y sólo cuando, el índice positivo (negativo) de inercia es igual a n.*

COROLARIO. *Una forma cuadrática es de signo constante cuando, y sólo cuando, uno de los índices de inercia es igual a cero.*

La ley de inercia permite dar cierta clasificación de las formas cuadráticas reales. Llamemos dos formas cuadráticas *equivalentes afines*, si para cada una de ellas se puede elegir tal base que las matrices de dichas formas cuadráticas se hagan iguales. En este caso diremos también que con ayuda de una transformación regular una forma cuadrática se convierte en la otra. Es fácil comprobar que la equivalencia afín de las formas cuadráticas es una relación de equivalencia y dos formas cuadráticas son equivalentes cuando, y sólo cuando, sus matrices en una misma base son congruentes. Por eso, de la ley de inercia proviene que todas las formas cuadráticas reales en el espacio lineal K_n pueden dividirse en clases disjuntas en cada una de las cuales entran formas cuadráticas equivalentes afines y sólo ellas. La clase se caracteriza por el rango y la signatura. La citada división en clases se denomina *clasificación afín* de las formas cuadráticas reales.

Para cualquier rango prefijado r de las formas cuadráticas en una clasificación dada siempre existen dos clases «extremas», las de signaturas $+r$ y $-r$. La primera clase consta de todas las formas cuadráticas no negativas de rango r. A la segunda clase pertenecen todas las formas cuadráticas no positivas de rango r. Ambas clases juntas contienen todas las formas cuadráticas de rango r de signo constante y sólo éstas.

La capacidad de una forma cuadrática de conservar su signo constante se establece, generalmente, por su reducción a la forma canónica, sirviéndose de uno de los métodos descritos. Sin embargo, en algunos casos son de gran interés los criterios inmediatos de la capacidad de conservar los signos constantes. Al tomar en consideración la gran importancia que tienen precisamente tales formas cuadráticas, realizaremos para éstas unas investigaciones adicionales, limitándonos, principalmente, a la consideración de las formas cuadráticas en un espacio real. Esta vez también consideraremos que la matriz de una forma cuadrática es simétrica real. Para el caso de un espacio complejo los resultados serán los mismos, mientras que las demostraciones se diferenciarán en pequeños detalles.

TEOREMA 94.2 (*Criterio de Sylvester*). *Para que una forma cuadrática sea definida positiva, es necesario y suficiente que todos los menores principales de la matriz de esta forma sean positivos.*

DEMOSTRACIÓN. NECESIDAD. Supongamos que una forma cuadrática con la matriz A es definida positiva. En este caso existe una transformación regular con la matriz P que reduce la forma a una suma de cuadrados. Conforme a (91.9), esto significa que $E = P'AP$

o bien $A = (P^{-1})' P^{-1}$. Haciendo uso de la fórmula Binet—Cauchy obtenemos

$$A\begin{pmatrix}1 & 2 \dots s\\ 1 & 2 \dots s\end{pmatrix} = \sum_{1\leqslant k_1<k_2<\dots<k_s\leqslant n} (P^{-1})' \times$$

$$\times \begin{pmatrix}1 & 2 & \dots & s\\ k_1 & k_2 & \dots & k_s\end{pmatrix} P^{-1}\begin{pmatrix}k_1 & k_2 \dots k_s\\ 1 & 2 \dots s\end{pmatrix} =$$

$$= \sum_{1\leqslant k_1<k_2<\dots k_s\leqslant n} \left(P^{-1}\begin{pmatrix}k_1 & k_2 \dots k_s\\ 1 & 2 \dots s\end{pmatrix}\right)^2.$$

Como la matriz P es regular, en las primeras s columnas hay al menos un vector distinto de cero. Por consiguiente, para todo s el segundo miembro de la igualdad obtenida es positivo.

SUFICIENCIA. Supongamos ahora que todos los menores principales de la matriz A de cierta forma cuadrática son positivos. Con ayuda de una transformación, determinada por las fórmulas (93.7), reduzcamos esta forma al tipo canónico. De acuerdo con la hipótesis del teorema y las fórmulas (93.4), todos los coeficientes del tipo canónico serán positivos, es decir, la forma cuadrática es definida positiva.

COROLARIO. *Para que una forma cuadrática sea definida negativa, es necesario y suficiente que todos los menores principales de orden impar sean negativos y todos los menores de orden par, positivos.*

La demostración se deduce del criterio de Sylvester y del hecho de que si A es una matriz de la forma cuadrática definida negativa, entonces $-A$ será una matriz de la forma cuadrática definida positiva.

TEOREMA 94.3 (*criterio de Jacobi*). *Para que una forma cuadrática sea definida positiva, es necesario y suficiente que todos los coeficientes del polinomio característico de la matriz de la forma sean distintos de cero y tengan signos alternados.*

DEMOSTRACION. NECESIDAD. Según se ha observado, por medio de una transformación de las variables con una matriz ortogonal, una forma cuadrática dada puede ser reducida al tipo canónico cuyos coeficientes serán los valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$ de la matriz de la forma. De acuerdo con las condiciones del teorema, los valores propios deben ser positivos. El polinomio característico $f(\lambda)$ es igual a

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots$$
$$\dots + a_1\lambda + a_0$$

siendo todos los coeficientes suyos no nulos con signos alternados, lo que se desprende directamente de las fórmulas de Viète para los coeficientes a_i.

SUFICIENCIA. Supongamos que los coeficientes del polinomio característico son distintos de cero y tienen signos alternados. Las raíces de este polinomio, como valores propios de la matriz simétrica, serán reales y nos resta por señalar que son positivos. Supondremos que esta afirmación está demostrada para todos los polinomios de grado $n-1$. Ya que todos los coeficientes $f'(\lambda)$ son distintos de cero y tienen signos alternados, entonces, por hipótesis, $f'(\lambda)$ tiene $n-1$ raíces positivas. Del curso del análisis matemático se conoce que si un polinomio sólo tiene raíces reales, éstas se dividen por las raíces de la derivada. Por eso $f(\lambda)$ tiene por lo menos $n-1$ raíces positivas. La última raíz será también positiva debido a que es positivo el producto de todas las raíces.

Los criterios para las formas cuadráticas no negativas y no positivas son mucho más complejos y esto se debe, en lo principal, a que en estos casos las matrices son degeneradas. Uno de los métodos que se usa para investigar el carácter constante de los signos de una forma cuadrática está vinculado con la reducción de su matriz a la forma simétrica (93.8), (93.9) como también con el estudio de esta última forma. En virtud de que las matrices de signos constantes están estrechamente enlazadas entre sí, nos limitaremos a la consideración de las matrices no negativas.

Llamemos a la matriz H *matriz de conmutaciones*, si cada una de sus filas y cada columna contienen sólo un elemento no nulo y todos los elementos no nulos son iguales a la unidad. Evidentemente, al multiplicar una matriz arbitraria A a la derecha por la matriz de conmutaciones H, en la matriz A se cambian entre sí sus columnas y al multiplicar a la izquierda, se cambian entre sí las filas.

LEMA 94.1. *Para una matriz regular arbitraria A existe tal matriz de conmutaciones H que en la matriz AH todos los menores principales son distintos de cero.*

DEMOSTRACION. A es una matriz regular. Por consiguiente, su primera fila contiene por lo menos un elemento no nulo. Al colocar la columna correspondiente en lugar de la primera, haremos no nulo el menor principal de primer orden. Supongamos que mediante la permutación de las columnas hemos logrado que todos los menores principales de orden hasta k no son iguales a cero. Ahora, si al permutar las últimas $n-k$ columnas no se puede conseguir que el menor principal de orden $(k+1)$ sea no nulo, esto significa que en las primeras $k+1$ filas de la matriz A no existe un menor no nulo de orden $k+1$, es decir, la matriz A debe ser degenerada. La contradicción con la hipótesis del lema significa que el último está demostrado.

TEOREMA 94.4. *Para que una forma cuadrática de rango r con la matriz A sea no negativa, es necesario y suficiente que exista una matriz de conmutaciones H tal, para la cual en la matriz $H'AH$ los primeros r menores principales sean positivos.*

DEMOSTRACIÓN. NECESIDAD. Supongamos que la forma cuadrática de rango r con la matriz A es no negativa. En este caso existe una matriz regular P tal que $A = (P^{-1})'E_rP^{-1}$, donde E_r es una matriz diagonal cuyos r primeros elementos son iguales a uno y los demás, a cero. De acuerdo con el lema 94.1, existe una matriz de conmutaciones H, para la cual todos los menores principales de la matriz $P^{-1}H$ son distintos de cero.

Haciendo uso de la fórmula de Binet—Cauchy, encontramos para $1 \leqslant s < r$

$$H'AH\begin{pmatrix}1 & 2 \dots s\\ 1 & 2 \dots s\end{pmatrix} =$$

$$= \sum_{1\leqslant k_1<k_2<\dots<k_s\leqslant n} (P^{-1}H)'\begin{pmatrix}1 & 2 & \dots & s\\ k_1 & k_2 & \dots & k_s\end{pmatrix} \times$$

$$\times (E_rP^{-1}H)\begin{pmatrix}k_1 & k_2 \dots & k_s\\ 1 & 2 \dots & s\end{pmatrix} =$$

$$= \sum_{1\leqslant k_1<k_2<\dots<k_s\leqslant r} \left\{(P^{-1}(H)\begin{pmatrix}k_1 & k_2 \dots & k_s\\ 1 & 2 \dots & s\end{pmatrix}\right\}^2 \geqslant$$

$$\geqslant \left\{(P^{-1}H)\begin{pmatrix}1 & 2 \dots s\\ 1 & 2 \dots s\end{pmatrix}\right\}^2 > 0.$$

SUFICIENCIA. Supongamos que para la forma cuadrática de rango r con la matriz A existe una matriz de conmutaciones H tal que en la matriz $H'AH$ los primeros r menores principales son positivos. La matriz $H'AH$ puede reducirse, de acuerdo con el teorema 93.1, al tipo canónico mediante la transformación con una matriz triangular. De conformidad con (93.4), los coeficientes no nulos del tipo canónico de la matriz $H'AH$ y, por tanto, de la matriz A serán positivos, es decir, la forma cuadrática es no negativa.

En cuanto a las formas cuadráticas de signo no constante, conviene decir que para ellas no existen análogos completos de los teoremas 94.1, 94.4. Sólo tiene lugar el

TEOREMA 94.5. *Si una forma cuadrática tiene una matriz simétrica A del tipo* (93.8), (93.9), *sus índices de inercia coinciden con los de inercia para la forma cuadrática «truncada» que se determina mediante la matriz B de* (93.8).

DEMOSTRACIÓN. De acuerdo con el teorema 93.1, la matriz A puede ser reducida a la forma canónica mediante la transformación con una matriz triangular derecha, con la particularidad de que para los coeficientes no nulos del tipo canónico se verifican las correlaciones (93.4). Pero la matriz B de la forma cuadrática «truncada» satisface también las condiciones del teorema 93.1 y para los coeficientes de su tipo canónico se verifican asimismo las correlaciones (93.4). Por eso, los índices de inercia de las formas cuadráticas, determinadas por las matrices A y B, coinciden.

El singular interés respecto a las formas cuadráticas de signo constante se debe al gran campo de sus aplicaciones. Una de las aplicaciones más importantes es la introducción de una métrica en un espacio lineal. Toda forma bilineal, polar respecto a cierta forma cuadrática definida positiva, puede considerarse como producto escalar y, por lo tanto, con su ayuda podemos convertir un espacio lineal en un espacio euclídeo o unitario. El cumplimiento de los axiomas para estos espacios es obvio. Una significación no menos importante para introducir una métrica, especialmente en los subespacios, tienen también las formas no negativas. Como ejemplo de la utilización del carácter constante de los signos demostremos que es válido el

TEOREMA 94.6. *Para que una forma bilineal hermitiana regular sea reducible a la forma diagonal, es suficiente que su parte simétrica (o antisimétrica) sea estrictamente de signo constante.*

DEMOSTRACION. Consideraremos el caso en que la parte simétrica es definida positiva. Sea A una matriz de la forma bilineal, entonces $\frac{1}{2}(A+A^*)$ será la matriz de la parte simétrica. Puesto que la parte simétrica es definida positiva, la matriz $\frac{1}{2}(A+A^*)$ es congruente de la matriz unidad según Hermite. Por consiguiente, existe una matriz regular S tal que

$$\frac{1}{2} S'(A+A^*)\overline{S} = E. \tag{94.7}$$

Mostremos que la matriz $S'A\overline{S}$ es normal. De (94.7) tenemos

$$\overline{S}S' = 2(A+A^*)^{-1}.$$

Por ello

$$\begin{aligned}(S'A\overline{S})(S'A\overline{S})^* - (S'A\overline{S})^*(S'A\overline{S}) &= S'(A\overline{S}S'A^* - A^*\overline{S}S'A) = \\ &= 2S'(A(A+A^*)^{-1}A^* - A^*(A+A^*)^{-1}A)\overline{S} = \\ &= 2S'((A^{*-1}(A+A^*)A^{-1})^{-1} - (A^{-1}(A+A^*)A^{*-1})^{-1})\overline{S} = \\ &= 2S'((A^{*-1}+A^{-1})^{-1} - (A^{*-1}+A^{-1})^{-1})\overline{S} = 0.\end{aligned}$$

Siendo normal, la matriz $S'AS$ se reduce a la forma diagonal mediante la transformación congruente según Hermite con una matriz unitaria.

De este modo, la matriz A de la forma bilineal hermitiana es congruente según Hermite de la matriz diagonal, lo que se trataba de demostrar. Todos los casos restantes se examinan de modo análogo.

Ejercicios.

1. Demuéstrese que si todos los menores principales de una matriz simétrica real o hermitiana compleja son distintos de cero, entonces el número de sus valores propios positivos y negativos coincide respectivamente con el número de los términos positivos y negativos de la sucesión (93.4).
2. Demuéstrese que si una matriz es definida positiva, todo menor diagonal es positivo.
3. Demuéstrese que una matriz simétrica de rango r siempre tiene un menor diagonal, aunque sea único, de rango r, distinto de cero.
4. Demuéstrese que el elemento máximo de una matriz definida positiva se encuentra en la diagonal principal.
5. Demuéstrese que la matriz A es definida positiva, si para todo i

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

6. Demuéstrese que para toda matriz simétrica A de rango r existe una matriz de conmutaciones H tal que entre los primeros r menores principales de la matriz $H'AH$ no hay dos adyacentes que sean nulos, con la particularidad de que el menor de r-ésimo orden es distinto de cero.
7. Demuéstrese que la matriz $H'AH$ del ejercicio 6 puede ser representada en la forma $H'AH = S'DS$, donde S es una matriz triangular derecha con elementos diagonales iguales a uno, D es una matriz diagonal celular cuyas células son de primero y segundo órdenes.
8. Demuéstrese que toda matriz no negativa de rango r puede ser representada como la suma de r matrices no negativas de rango 1.
9. Sean A, B las matrices definidas positivas con los elementos a_{ij}, b_{ij}. Demuéstrese que la matriz C con los elementos $c_{ij} = a_{ij}b_{ij}$ es también definida positiva

§ 95. Hipersuperficies de segundo grado

Con el estudio de las formas cuadráticas reales está estrechamente relacionada la investigación de otros objetos, a saber, hipersuperficies de segundo grado. Deseando subrayar el carácter geométrico de muchas propiedades de las hipersuperficies, en adelante llamaremos a los vectores casi siempre puntos del espacio $\mathbf{R}_n$.

Se denomina *hipersuperficie f de segundo grado* en el espacio $\mathbf{R}_n$ un conjunto de puntos cuyas coordenadas $x_1, x_2, \ldots, x_n$ satisfacen la ecuación

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} x_i x_j - 2 \sum_{k=1}^{n} b_k x_k + c = 0, \tag{95.1}$$

donde a_{ji}, b_k, c son unos números reales.

Simplifiquemos la anotación. Al igual que en el caso de las formas cuadráticas, supondremos que la matriz A con los coeficientes a_{ji} es simétrica. Designemos con b un vector con las coordenadas $b_1, b_2, \ldots, b_n$. Introduzcamos en el espacio $\mathbf{R}_n$ un producto escalar como suma de productos de las coordenadas, tomados dos a dos.

Ahora, la hipersuperficie f de segundo grado en el espacio $\mathbf{R}_n$ puede considerarse como un conjunto de los puntos x del espacio euclídeo $\mathbf{R}_n$ que satisfacen la ecuación

$$(Ax, x) - 2(b, x) + c = 0 \tag{95.2}$$

o bien, por ser la matriz A simétrica, la ecuación

$$(x, Ax) - 2(b, x) + c = 0.$$

La investigación de las hipersuperficies de segundo grado la empezaremos con el estudio de la disposición conjunta de estas superficies y de líneas rectas. Tomemos una recta arbitraria en el espacio $\mathbf{R}_n$. Supongamos que esta recta pasa por el punto x_0 y tiene un vector director l. Los puntos x de dicha recta se definen mediante la igualdad

$$x = x_0 + lt \tag{95.3}$$

para cualesquiera números reales t. Al sustituir la expresión dada para x en (95.2), obtendremos

$$t^2 (Al, l) - 2t((b, l) - (Al, x_0)) + \\ + (Ax_0, x_0) - 2(b, x_0) + c = 0. \tag{95.4}$$

De este modo, los puntos de intersección de la recta (95.3) con la hipersuperficie (95.2) se determinan por las raíces de la ecuación cuadrática (95.4).

Diremos que la recta (95.3) con el vector director l es de dirección *no asintótica* (*asintótica*) respecto de la hipersuperficie (95.2), si $(Al, l) \neq 0$ $((Al, l) = 0)$.

Consideraremos una recta cualquiera que tiene la dirección no asintótica l y atraviesa la hipersuperficie. Los puntos de intersección determinan en cada una de estas rectas un segmento al cual llamaremos, por analogía con la geometría elemental, *cuerda*. Designemos con L el conjunto de puntos medios de todas las cuerdas. Si los extremos de una cuerda son contraídos a un punto, éste se considerará también como punto medio de la cuerda. Probemos que L pertenece a cierta hipersuperficie.

Los extremos de cualquier cuerda se determinan por los valores del parámetro t coincidentes con las raíces de la ecuación (95.4). Por ello el punto medio de la cuerda se determina por el valor de t igual a la semisuma de las raíces. De confirmidad con las fórmulas de Viète esto nos da

$$t = \frac{(b, l) - (Al, x_0)}{(Al, l)}. \tag{95.5}$$

Si z_0 es el punto medio de la cuerda, entonces

$$z_0 = x_0 + \frac{(b, l) - (Al, x_0)}{(Al, l)} l.$$

Ahora tenemos

$$(Al, z_0) = \left(Al,\ x_0 + \frac{(b, l) - (Al, x_0)}{(Al, l)}\, l \right) =$$
$$= (Al, x_0) + \frac{(b, l) - (Al, x_0)}{(Al, l)} (Al, l) = (b, l).$$

Así pues, los puntos medios de todas las cuerdas satisfacen la ecuación

$$(Al, x) = (b, l). \tag{95.6}$$

Como el segundo miembro de la ecuación no depende de x_0, entonces, de acuerdo con la fórmula (46.8), esta ecuación determina un hiperplano cuyo vector normal es igual a Al.

El hiperplano (95.6) se llama *hiperplano diametral* conjugado de la dirección l respecto a la hipersuperficie (95.2).

La forma explícita de la ecuación del hiperplano diametral permite establecer toda una serie de propiedades importantes que poseen las hipersuperficies de segundo grado. Sea A una matriz regular. Entonces, para cualesquiera vectores linealmente independientes $l_1, l_2, \ldots, l_n$ serán también linealmente independientes los vectores $Al_1, Al_2, \ldots, Al_n$. Supondremos luego que todas las direcciones $l_1, l_2, \ldots, l_n$ son no asintóticas. Esto tendrá lugar a ciencia cierta en el caso, por ejemplo, cuando la forma cuadrática (Ax, x) es definida positiva. Por consiguiente, se puede construir un sistema de n hiperplanos diametrales conjugados de las direcciones $l_1, l_2, \ldots, l_n$. Los hiperplanos tendrán un único punto común x^*. Ahora, de la fórmula (95.6) se desprenden las igualdades

$$(Ax^* - b, l_i) = 0$$

para $i = 1, 2, \ldots, n$. En virtud de la independencia lineal de los vectores l_i, esto significa que $Ax^* - b = 0$, es decir, el punto x^* no es otra cosa que la solución del sistema de ecuaciones algebraicas lineales

$$Ax = b. \tag{95.7}$$

La solución del sistema con una matriz regular es única, razón por la cual el punto construido x^* no depende, en realidad, de cómo se escogen los vectores $l_1, l_2, \ldots, l_n$.

Unos cálculos simples muestran que para todo punto x^* es válida la correlación

$$(Ax, x) - 2(b, x) + c = (A(x - x^*), x - x^*) +$$
$$+ 2(Ax^* - b, x - x^*) + (Ax^*, x^*) - 2(b, x^*) + c. \tag{95.8}$$

Si, en cambio, x^* es la solución del sistema (95.7), entonces respecto de tal punto la hipersuperficie (95.2) posee la propiedad importante de simetría. A saber, cualquiera que sea x, el primer miembro de (95.2) toma valores iguales en los puntos

$$x = x^* + (x - x^*), \qquad x' = x^* - (x - x^*). \tag{95.9}$$

De aquí se deduce, en particular, que ambos puntos, x, x', se ubican o no se ubican en la hipersuperficie (95.2) simultáneamente. La igualdad

$$x^* = \frac{1}{2}(x + x')$$

permite llamar al punto x^* *centro de simetría* de la hipersuperficie. Si en la hipersuperficie (94.2) se dispone aunque sea un solo punto de $\mathbf{R}_n$, el centro de simetría se denomina *real*. En el caso contrario se llama *imaginario*.

Sea, ahora, x^* un centro de simetría, es decir, para todo x el primer miembro de (95.2) toma valores iguales en los puntos x, x'. Por consiguiente,

$$(Ax, x) - 2(b, x) + c = (Ax', x') - 2(b, x') + c.$$

De acuerdo con (95.8), (95.9), esto es posible sólo en el caso en que para cualquier x

$$(Ax^* - b, x - x^*) \equiv 0.$$

Mas, la última identidad es válida cuando, y sólo cuando, $Ax^* - b = 0$, es decir, cuando el punto x^* sea la solución del sistema (95.7). Cabe señalar que aquí nunca suponíamos la regularidad de la matriz A, ni tampoco la presencia de otras sus singularidades, salvo la simetría. Por esta razón:

Para que el sistema $Ax = b$ tenga solución, es necesario y suficiente que la hipersuperficie (95.2) cuente con un centro de simetría. El conjunto de todas las soluciones coincide con el conjunto de todos los centros de simetría.

De este modo, se pone de manifiesto una relación muy profunda entre los sistemas de ecuaciones algebraicas lineales y las hipersuperficies de segundo grado. Esta relación se utiliza ampliamente al construir los más diversos algoritmos de cálculo. En particular, en la construcción del sistema de hipersuperficies diametrales se basa un gran grupo de métodos que forma parte del grupo de los llamados métodos de direcciones conjugadas. Estos métodos se tratarán en el último capítulo de la obra.

En el caso general la investigación de las hipersuperficies de segundo grado puede fundarse en la reducción de ellas a la forma canónica, casi por analogía completa con las formas cuadráticas. Pero en este caso, además de las transformaciones regulares lineales de las variables, se exigirán las operaciones de desplazamiento.

Consideremos una transformación cualquiera de las variables $x = Py$ que reduce la forma cuadrática (Ax, x) a la forma normal. En términos de las variables $y_1, y_2, \ldots, y_n$ la ecuación de la hipersuperficie tendrá por expresión

$$y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 -$$
$$- 2d_1y_1 - \ldots - 2d_ry_r - 2d_{r+1}y_{r+1} - \ldots - 2d_ny_n + c = 0.$$

Realicemos ahora el desplazamiento de las variables de acuerdo con las fórmulas

$$z_i = \begin{cases} y_i - d_i, & 1 \leqslant i \leqslant k, \\ y_i + d_i, & k+1 \leqslant i \leqslant r, \\ y_i & r+1 \leqslant i \leqslant n. \end{cases}$$

En términos de estas variables la ecuación toma la forma

$$z_1^2 + \ldots + z_k^2 - z_{k+1}^2 - \ldots - z_r^2 - 2d_{r+1}z_{r+1} - \ldots - 2d_nz_n + p = 0.$$

Supongamos que uno de los números $d_{r+1}, \ldots, d_n$, por ejemplo d_n, es distinto de cero. Hagamos

$$v_i = \begin{cases} z_i, & i < n, \\ d_{r+1}z_{r+1} + \ldots + d_nz_n, & i = n, \end{cases}$$

y a continuación realicemos nuevamente un desplazamiento

$$u_i = \begin{cases} v_i, & i < n \\ v_i - \frac{p}{2} & i = n. \end{cases}$$

Ahora, la ecuación de la hipersuperficie adquiere la forma:

$$\pm u_1^2 \pm u_2^2 \pm \ldots \pm u_r^2 - 2u_n = 0, \quad 1 \leqslant r \leqslant n-1. \qquad (95.10)$$

Si entre los números $d_{r+1}, \ldots, d_n, p$ no hay ninguno que sea igual a cero, la ecuación de la hipersuperficie toma la forma siguiente:

$$\pm u_1^2 \pm u_2^2 \pm \ldots \pm u_r^2 = 0, \qquad 1 \leqslant r \leqslant n. \qquad (95.11)$$

Y, por fin, si los números $d_{r+1}, \ldots, d_n$ son nulos, mientras que $p \neq 0$, entonces, al poner $u_i = z_i/|p|^{1/2}$ para todo i, obtendremos una forma más para la ecuación de la hipersuperficie. A saber,

$$\pm u_1^2 \pm u_2^2 \pm \ldots \pm u_r^2 \pm 1 = 0, \qquad 1 \leqslant r \leqslant n. \qquad (95.12)$$

En virtud de la ley de inercia de las formas cuadráticas, las superficies definidas por diversas ecuaciones del tipo (95.10)—(95.12) no pueden convertirse la una en la otra con ayuda de la transformación lineal de las variables y un desplazamiento. En este caso deben considerarse diferentes las ecuaciones que no pueden ser transformadas la una en la otra multiplicándolas por (-1) y cambiando la numeración de las coordenadas. Igual que en el caso de las formas cuadráticas, esta vez obtuvimos también una partición de todas las hipersuperficies de segundo grado en clases disjuntas.

Al reducir las hipersuperficies de segundo grado a la forma canónica se utilizan con frecuencia sólo las operaciones de traslado y transformaciones lineales de las variables con matrices ortogonales. Esto se debe, principalmente, a que ambos tipos de las transformaciones indicadas no cambian las distancias entre los puntos. En este caso las formas canónicas serán un tanto diferentes, aunque en total se obtienen del mismo modo que las anteriores. Por ejemplo, en el caso del espacio $\mathbf{R}_2$, la hipersuperficie de segundo grado puede ser reducida sólo a uno de los siguientes tipos:

$$
\begin{aligned}
&\text{I. } \lambda_1 x^2 + \lambda_2 y^2 + a_0 = 0,\\
&\text{II. } \mu_2 y^2 + b_0 x = 0,\\
&\text{III. } \lambda_1 x^2 + a_0 = 0,
\end{aligned}
\tag{95.13}
$$

y en el caso del espacio R_3, a uno de los siguientes tipos:

$$
\begin{aligned}
&\text{I. } \lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + a_0 = 0,\\
&\text{II. } \lambda_1 x^2 + \lambda_2 y^2 + b_0 z = 0,\\
&\text{III. } \lambda_1 x^2 + \lambda_2 y^2 + a_0 = 0,\\
&\text{IV. } \lambda_1 y^2 + b_0 x = 0,\\
&\text{V. } \lambda_1 x^2 + a_0 = 0.
\end{aligned}
\tag{95.14}
$$

En todas las ecuaciones (95.13), (95.14) los coeficientes de las variables escritas son distintos de cero. El término independiente puede ser igual a cero. De acuerdo con la terminología aceptada, las hipersuperficies en el espacio $\mathbf{R}_2$ se llamarán *líneas de segundo orden* y en el espacio $\mathbf{R}_3$, *superficies de segundo grado*. Tomando en consideración los intereses de los diferentes apartados de las matemáticas, estudiaremos más detalladamente las líneas y las superficies de segundo grado según sus formas canónicas (95.13), (95.14).

Ejercicios.

1. Sea A una matriz definida positiva. Demuéstrese que en la solución del sistema $Ax = b$ la expresión $(Ax, x) - 2(b, x)$ alcanza su valor mínimo.

2. Sea A una matriz definida positiva. Demuéstrese que en la recta (95.3) la expresión $(Ax, x) - 2(b, x)$ alcanza su valor mínimo para el valor de t tomado de (95.5).

3. Demuéstrese que para que una dirección cualquiera sea no asintótica para la hipersuperficie (95.2), es necesario y suficiente que la forma cuadrática (Ax, x) sea definida positiva o negativa.

4. ¿Qué propiedad de simetría posee un hiperplano diametral conjugado de la dirección l, si l es un vector propio de la matriz A, correspondiente al valor propio no nulo?

5. Demuéstrese que el sistema $Ax = b$ no tiene solución cuando, y sólo cuando, el hiperplano (95.2) se reduce a la forma canónica (95.10).

§ 96. Líneas de segundo orden

Estudiaremos las líneas de segundo orden mediante las ecuaciones (95.13). Supongamos que la ecuación de una línea tiene por expresión

$$\lambda_1 x^2 + \lambda_2 y^2 + a_0 = 0. \tag{96.1}$$

1.1. *El número a_0 no es nulo; los números λ_1, λ_2, son de signos iguales, contrarios al signo de a_0.* Escribamos (96.1) en la forma siguiente

$$\frac{x^2}{\left(\sqrt{-\frac{a_0}{\lambda_1}}\right)^2} + \frac{y^2}{\left(\sqrt{-\frac{a_0}{\lambda_2}}\right)^2} = 1$$

y designemos

$$a = \sqrt{-\frac{a_0}{\lambda_1}}, \quad b = \sqrt{-\frac{a_0}{\lambda_2}} \tag{96.2}$$

Según la condición, los números a y b son reales, por lo cual la ecuación (96.1) es equivalente a la ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{96.3}$$

Una línea, descrita por esta ecuación, se denomina *elipse* (fig. 96.1) y la propia ecuación se llama *ecuación canónica de la elipse.* Demos a conocer algunas propiedades de la elipse. Una elipse es una línea acotada. Como se deduce de la ecuación (96.3), para todos los puntos de la elipse se tiene: $|x| \leqslant a$, $|y| \leqslant b$. La elipse cuenta con dos ejes de simetría: el eje Ox y el eje Oy, como también un centro de simetría que es el origen de coordenadas. Esto se deduce del hecho que a la par con un punto de coordenadas (x, y), a la elipse pertenecen los puntos que tienen las coordenadas $(x, -y)$, $(-x, y)$, $(-x, -y)$. Los ejes de simetría se llaman *ejes principales* de la elipse; el centro de simetría es a la vez *centro* de la elipse. Si $a > b$, entonces Ox lleva el nombre de eje *mayor* de la elipse y Oy, eje *menor* de la elipse. Los puntos de intersección de los ejes principales de la elipse con la propia elipse se llaman *vértices* de

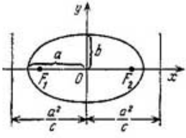

Fig. 96.1.

la elipse. Cuando $a = b$, la elipse se convierte en una circunferencia de radio a y centro en el origen de coordenadas. Supongamos, para concretar, que $a > b$ y designemos

$$c^2 = a^2 - b^2. \tag{96.4}$$

Los puntos F_1, F_2 con las coordenadas $(-c, 0)$, $(+c, 0)$ se denominan *focos* de la elipse.

TEOREMA 96.1. *La suma de las distancias desde cualquier punto de la elipse hasta sus focos es una magnitud constante, igual a $2a$.*

DEMOSTRACION. Para todo punto $M(x, y)$ de la elipse tenemos

$$y^2 = b^2 - \frac{b^2x^2}{a^2}.$$

Para este mismo punto calculamos

$$\rho(M, F_2) = ((x-c)^2 + y^2)^{1/2} = \left(x^2 - 2xc + c^2 + b^2 - \frac{b^2x^2}{a^2}\right)^{1/2} =$$
$$= \left(x^2\left(1 - \frac{b^2}{a^2}\right) - 2xc + a^2\right)^{1/2} = \left(\frac{x^2c^2}{a^2} - 2xc + a^2\right)^{1/2} =$$
$$= \left(\left(-\frac{c}{a}x + a\right)^2\right)^{1/2} = -\frac{c}{a}x + a.$$

La última igualdad es válida, puesto que $-\frac{c}{a}x + a > 0$ porque $|x| \leqslant a$ y $c/a < 1$. Luego,

$$\rho(M, F_1) = ((x+c)^2 + y^2)^{1/2} = \left(x^2 + 2xc + c^2 + b^2 - \frac{b^2x^2}{a^2}\right)^{1/2} =$$
$$= \left(x^2\left(1 - \frac{b^2}{a^2}\right) + 2xc + a^2\right)^{1/2} = \left(\frac{x^2c^2}{a^2} + 2xc + a^2\right)^{1/2} =$$
$$= \left(\left(\frac{cx}{a} + a\right)^2\right)^{1/2} = \frac{c}{a}x + a.$$

En definitiva tenemos

$$\rho(M, F_1) + \rho(M, F_2) = -\frac{c}{a}x + a + \frac{c}{a}x + a = 2a.$$

I.2. *El número a_0 no es nulo; los números λ_1, λ_2, a_0 son de signos iguales.* Designemos

$$a = \sqrt{\frac{a_0}{\lambda_1}}, \quad b = \sqrt{\frac{a_0}{\lambda_2}}, \tag{96.5}$$

entonces la ecuación (96.1) es equivalente a la ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1. \tag{96.6}$$

Está claro que no hay ningún punto del plano que satisfaga (96.6). La ecuación (96.6) suele tratarse como la ecuación de una elipse *imaginaria*.

I.3. *El número a_0 es nulo; los números λ_1, λ_2 son de signos iguales.* Designemos

$$a=\sqrt{\frac{1}{|\lambda_1|}}, \qquad b=\sqrt{\frac{1}{|\lambda_2|}},$$

en este caso la ecuación (96.1) es equivalente a la siguiente:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=0. \tag{96.7}$$

Está claro que solamente el origen de coordenadas satisface la ecuación (96.7). La ecuación (96.7) suele tratarse como la ecuación de una elipse *degenerada*.

I.4. *El número a_0 no es igual a cero; los números λ_1, λ_2 son de signos contrarios.* Por introducción de los nuevos coeficientes análogos a (96.2), (96.5), la ecuación (96.1) se reduce, salvo la redesignación de las variables, a una ecuación equivalente

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{96.8}$$

Una línea, descrita por esta ecuación, se llama *hipérbola* (fig. 96.2) y la propia ecuación, ecuación *canónica* de la hipérbola. A diferencia de la elipse, la hipérbola es una línea no acotada. Igual que en el caso de una elipse, los ejes de simetría de la hipérbola son los ejes de coordenadas y el centro de simetría es el origen de coordenadas. Los ejes de simetría se denominan *ejes principales* de la hipérbola; el centro de simetría, *centro* de la hipérbola. Uno de los ejes principales (Ox) se interseca con la hipérbola en dos puntos llamados *vértices* de la hipérbola. Este eje se llama eje *real* de la hipérbola. El otro eje (Oy) no tiene puntos comunes con la hipérbola y se llama, por eso, eje *imaginario* de la hipérbola.

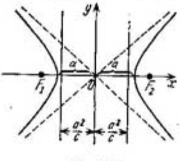

Fig. 96.2.

Designemos

$$c^2=a^2+b^2.$$

Los puntos F_1, F_2, cuyas coordenadas son $(-c, 0)$, $(+c, 0)$ se denominan *focos* de la hipérbola.

TEOREMA 96.2. *La magnitud absoluta de la diferencia entre las distancias desde cualquier punto de la hipérbola hasta sus focos es constante e igual a $2a$.*

DEMOSTRACION. Para todo punto $M(x, y)$ de la hipérbola se tiene

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

Para el mismo punto calculamos

$$\rho(M, F_2) = ((x-c)^2+y^2)^{1/2} = \left(x^2-2xc+c^2-b^2+\frac{b^2x^2}{a^2}\right)^{1/2} =$$
$$= \left(x^2\left(1+\frac{b^2}{a^2}\right)-2xc+a^2\right)^{1/2} = \left(\frac{x^2c^2}{a^2}-2xc+a^2\right)^{1/2} =$$
$$= \left(\left(\frac{c}{a}x-a\right)^2\right)^{1/2} = \left|\frac{c}{a}x-a\right|.$$

Luego

$$\rho(M, F_1) = ((x+c)^2+y^2)^{1/2} =$$
$$= \left(x^2+2xc+c^2+\frac{b^2}{a^2}x^2-b^2\right)^{1/2} = \left(x^2\left(1+\frac{b^2}{a^2}\right)+2xc+a^2\right)^{1/2} =$$
$$= \left(\frac{x^2c^2}{a^2}+2xc+a^2\right)^{1/2} = \left(\left(\frac{c}{a}x+a\right)^2\right)^{1/2} = \left|\frac{c}{a}x+a\right|.$$

Para todos los puntos de la hipérbola tenemos $|x| \geqslant a$ y $c/a > 1$. Por esto

$$\rho(M, F_2) = \begin{cases} \frac{c}{a}x-a & \text{para todo } x>0, \\ -\frac{c}{a}x+a & \text{para todo } x<0, \end{cases}$$

$$\rho(M, F_1) = \begin{cases} \frac{c}{a}x+a & \text{para todo } x>0, \\ -\frac{c}{a}x-a & \text{para todo } x<0. \end{cases}$$

En defenitiva,

$$|\rho(M, F_1) - \rho(M, F_2)| = 2a.$$

Consideraremos la parte de la hipérbola dispuesta en el primer cuadrante. Para esta parte $x \geqslant a$, $y \geqslant 0$. La ecuación (96.8) en el primer cuadrante es equivalente a la siguiente

$$y = \frac{b}{a}\sqrt{x^2-a^2},$$

si, naturalmente, se considera que $b > 0$, $a > 0$. Es fácil convencerse de que esta función puede ser representada en la forma siguiente:

$$y = \frac{b}{a}x - \frac{ba}{x+\sqrt{x^2-a^2}}. \tag{96.9}$$

A la par con la función (96.9) examinemos la ecuación de una recta

$$y' = \frac{b}{a}x. \tag{96.10}$$

Designaremos mediante $M(x, y)$ y $M'(x, y')$ los puntos de la hipérbola (96.9) y de la recta (96.10) que tienen una misma abscisa x.

Al crecer ilimitadamente x, la diferencia

$$y'-y=\frac{ba}{x+\sqrt{x^2-a^2}}$$

va decreciendo de manera monótona, quedando positiva, y tiende hacia cero. Por consiguiente, los puntos M y M' van acercándose, pero el punto M de la hipérbola queda siempre por debajo del punto M' en la recta (96.10).

Una propiedad análoga tiene lugar también para las otras partes de la hipérbola. Una de las rectas

$$y=\frac{b}{a}x, \qquad y=-\frac{b}{a}x \tag{96.11}$$

desempeña el papel de la recta (96.10). Estas rectas se llaman *asíntotas* de la hipérbola.

Observemos que hemos tratado (96.8) como la ecuación de la hipérbola. Sin embargo, del curso escolar se conoce otra ecuación que también se denomina *ecuación de la hipérbola*.

Teniendo presente (96.11), realicemos el siguiente cambio de coordenadas

$$x'=\frac{x}{a}-\frac{y}{b}, \qquad y'=\frac{x}{a}+\frac{y}{b}.$$

De (96.8) tenemos

$$\left(\frac{x}{a}-\frac{y}{b}\right)\left(\frac{x}{a}+\frac{y}{b}\right)=1.$$

Por lo tanto, en el nuevo sistema de coordenadas (no rectangular, en el caso general) la ecuación de la hipérbola tiene la forma

$$x'y'=1 \tag{96.12}$$

o bien

$$y'=\frac{1}{x'}.$$

Ésta es precisamente la ecuación conocida del curso escolar. La ecuación (96.12) se denomina ecuación de la hipérbola respecto de sus asíntotas.

I.5. *El número a_0 es igual a cero; los números λ_1, λ_2 son de signos contrarios.* Al efectuar el cambio habitual de los coeficientes, obtendremos la ecuación

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=0 \tag{96.13}$$

equivalente a la ecuación (96.1). De esta ecuación obtenemos

$$\left(\frac{x}{a}-\frac{y}{b}\right)\left(\frac{x}{a}+\frac{y}{b}\right)=0$$

o bien

$$y=\frac{b}{a}x, \qquad y=-\frac{b}{a}x. \tag{96.14}$$

De este modo, la ecuación (96.13) es la ecuación de una línea que se descompone en *dos rectas* (96.14) *que se cortan.*

Consideraremos ahora la segunda ecuación de (95.13). Ésta tiene la forma

$$\lambda_2 y^2 + b_0 x = 0. \tag{96.15}$$

II.6. *Ambos números* λ_2, b_0 *son distintos de cero.* Designemos

$$2p=-\frac{b_0}{\lambda_2}\neq 0.$$

Ahora, la ecuación (96.15) resulta equivalente a la siguiente:

$$y^2 = 2px. \tag{96.16}$$

La línea descrita por esta ecuación se denomina *parábola* (fig. 96.3) y la propia ecuación lleva el nombre de ecuación *canónica* de la parábola. Sin restringir la generalidad, se puede considerar que $p>0$, puesto que para $p<0$ se obtiene una línea simétrica respecto del eje Oy. Análogamente a la hipérbola, la parábola es una línea no acotada. Tiene solamente un eje de simetría, el eje Ox, y no tiene centro de simetría. El punto de intersección del eje de la parábola con la misma parábola se llama *vértice* de la parabola. El punto F cuyas coordenadas son $\left(\frac{p}{2}, 0\right)$ se denomina *foco* de la parábola. La recta L definida por la ecuación

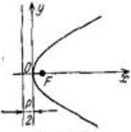

Fig. 96.3.

$$x=-\frac{p}{2} \tag{96.17}$$

se llama directriz de la parabola.

TEOREMA 96.3. *La distancia entre cualquier punto de una parábola y la directriz es igual a la distancia entre dicho punto y el foco de la parábola.*

DEMOSTRACION. Para cualquier punto $M(x, y)$ de la parábola tenemos

$$\rho(L, M)=x+\frac{p}{2},$$

y luego

$$\rho(F, M) = \left(\left(x-\frac{p}{2}\right)^2 + y^2\right)^{1/2} = \left(\left(x-\frac{p}{2}\right)^2 + 2px\right)^{1/2} =$$
$$= \left(x^2 - px + \frac{p^2}{4} + 2px\right)^{1/2} = \left(x^2 + px + \frac{p^2}{4}\right)^{1/2} =$$
$$= \left(\left(x+\frac{p}{2}\right)^2\right)^{1/2} = x + \frac{p}{2}$$

puesto que $x \geqslant 0$ y $p > 0$.

Consideraremos, por fin, la tercera ecuación de (95.13). Es de la forma más sencilla:

$$\lambda_1 x^2 + a_0 = 0. \qquad (96.18)$$

III.7. *El número a_0 no es igual a cero; el signo del número λ_1 es contrario al de a_0.* Designemos

$$a^2 = -\frac{a_0}{\lambda_1},$$

entonces la ecuación de la línea (96.18) será equivalente a la ecuación

$$x^2 - a^2 = 0 \qquad (96.19)$$

o bien

$$x = a, \qquad x = -a. \qquad (96.20)$$

Por consiguiente, la ecuación de la línea (96.19) es la ecuación de una línea que se descompone en dos rectas *paralelas* (96.20).

III.8. *El número a_0 no es igual a cero; el signo del número λ_1 coincide con el de a_0.* Designemos

$$a^2 = \frac{a_0}{\lambda_1}.$$

entonces la ecuación (96.18) será equivalente a la ecuación

$$x^2 + a^2 = 0. \qquad (96.21)$$

Está claro que no existe ningún punto del plano cuyas coordenadas satisfagan esta ecuación. De (96.21) suele decirse como de una ecuación que define *dos rectas imaginarias.*

III.9. *El número a_0 es igual a cero.* En este caso (96.18) es equivalente a la ecuación

$$x^2 = 0. \qquad (96.22)$$

Por analogía con la ecuación (96.19), suele decirse que la ecuación (96.22) define *dos rectas coincidentes*, cada una de las cuales se determina mediante la ecuación

$$x = 0.$$

Ha de señalarse que para todos los puntos de una elipse o una hipérbola tienen lugar las siguientes igualdades:

$$\rho(M, F_2) = \frac{c}{a}\left|x - \frac{a^2}{c}\right|, \tag{96.23}$$

$$\rho(M, F_1) = \frac{c}{a}\left|x + \frac{a^2}{c}\right|.$$

Las rectas α_i $(i = 1, 2)$, definidas por las ecuaciones

$$x - \frac{a^2}{c} = 0, \qquad x + \frac{a^2}{c} = 0, \tag{96.24}$$

se llaman *directrices* de la elipse y de la hipérbola. Atribuiremos a la directriz y al foco números idénticos, si se disponen en un mismo semiplano definido por el eje Oy. Ahora podemos mostrar que:

La razón entre las distancias $\rho(M, F_i)$ y $\rho(M, \alpha_i)$ es una magnitud constante para todos los puntos M de la elipse, hipérbola y parábola.

Para la parábola esta afirmación se deduce del teorema 96.3. Para la elipse y la hipérbola, de las ecuaciones (96.23), (96.24). La razón

$$e = \frac{\rho(M, F_i)}{\rho(M, \alpha_i)}$$

se denomina *excentricidad*. Se tiene

para la elipse:

$$e = \frac{c}{a} = \left(1 - \frac{b^2}{a^2}\right)^{1/2} < 1,$$

para la hipérbola:

$$e = \frac{c}{a} = \left(1 + \frac{b^2}{a^2}\right)^{1/2} > 1,$$

para la parábola:

$$e = 1.$$

Ejercicios.

1. ¿Qué representa en sí un hiperplano diametral conjugado de una dirección dada para las líneas de segundo orden?
2. Escríbanse las ecuaciones de una línea tangente para una elipse, una hipérbola y una parábola.
3. Demuéstrese que un rayo de luz, que sale de un foco de la elipse y se refleja de la tangente, pasa por el segundo foco.
4. Demuéstrese que un rayo de luz que sale del foco de una parábola y se refleja de la tangente, pasa paralelamente al eje de la parábola.
5. Demuéstrese que un rayo de luz que sale de un foco de la hipérbola y se refleja de la tangente, aparece saliente del segundo foco.

§ 97. Superficies de segundo grado

Pasemos ahora al estudio de las superficies de segundo grado, dadas en forma de las ecuaciones (95.14). Consideraremos primero la ecuación

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + a_0 = 0. \tag{97.1}$$

I.1. *El número a_0 es distinto de cero; los signos de todos los números λ_1, λ_2, λ_3 son idénticos y contrarios al signo de a_0.* El cambio habitual de los coeficientes nos da

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{97.2}$$

La superficie descrita por esta ecuación se llama *elipsoide* (fig. 97.1) y la propia ecuación (97.2), ecuación *canónica* del elipsoide. De la ecuación (97.2) se deduce que los planos de coordenadas son *planos de simetría* y el origen de coordenadas, *el centro de simetría*. Los números a, b, c se denominan *semiejes* del elipsoide. Un elipsoide es una superficie limitada, encerrada dentro de un paralelepípedo $|x| \leqslant a$, $|y| \leqslant b$, $|z| \leqslant c$. La línea de intersección del elipsoide con cualquier plano representa una elipse. En efecto, tal línea de intersección es una línea de segundo grado. Por ser el elipsoide una superficie limitada, esta línea también será limitada, pero la única línea limitada de segundo orden es una elipse.

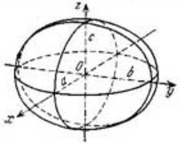

Fig. 97.1.

I.2. *El número a_0 es distinto de cero; los signos de todos los números λ_1, λ_2, λ_3, a_0 son idénticos.* La sustitución habitual de los coeficientes da

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1. \tag{97.3}$$

No hay ningún punto del espacio cuyas coordenadas satisfagan esta ecuación. De (97.3) suele decirse como de una ecuación de un elipsoide *imaginario.*

I.3. *El número a_0 es igual a cero; los signos de todos los números λ_1, λ_2, λ_3 son idénticos.* Tenemos

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0. \tag{97.4}$$

Esta ecuación se satisface sólo por el origen de coordenadas. Suele decirse que (97.4) es la ecuación de *una elipse degenerada.*

I.4. *El número a_0 no es igual a cero; los signos de λ_1, λ_2 coinciden y son contrarios a los de λ_3, a_0.* La sustitución habitual de los coeficientes nos da

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1. \qquad (97.5)$$

La superficie descrita por esta ecuación se llama ***hiperboloide de una hoja*** (fig. 97.2) y la misma ecuación, ecuación ***canónica*** del hiperboloide de una hoja. De la ecuación (97.5) se deduce que los planos de coordenadas son *los de simetría* y el origen de coordenadas es *el centro de simetría.* Examinemos las líneas L_h de intersección del hiperboloide de una hoja con los planos $z=h$. La ecuación de la proyección de tal línea sobre el plano Oxy se obtiene de la ecuación (97.5), si ponemos en ésta $z=h$. Es fácil ver que esta línea es una elipse

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

donde

$$a^*=a\sqrt{1+h^2/c^2},$$
$$b^*=b\sqrt{1+h^2/c^2},$$

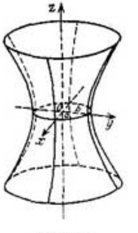

Fig. 97.2.

con la particularidad de que sus dimensiones crecen ilimitadamente cuando $h\to+\infty$. Las secciones que se obtienen al cortar el hiperboloide de una hoja por los planos Oyz y Oxz representan en sí las hipérbolas.

De este modo, el hiperboloide de una hoja representa en sí una superficie compuesta por una hoja y semejante a un tubo. Dicha superficie se extiende ilimitadamente en las direcciones positiva y negativa del eje Oz.

I.5. *El número a_0 no es igual a cero; los signos de λ_1, λ_2, a_0 coinciden y son contrarios al de λ_3.* Por analogía con (97.5), tenemos

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1. \qquad (97.6)$$

La superficie descrita por esta ecuación se denomina *hiperboloide de dos hojas* (fig. 97.3) y la propia ecuación, ecuación *canónica* del hiperboloide de dos hojas. Los planos de coordenadas son *planos de simetría* y el origen de coordenadas es *el centro de simetría.* Las líneas de intersección L_h del hiperboloide de dos hojas con los planos $z=h$ representan elipses cuyas proyecciones sobre el plano Oxy tienen por expresión

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

donde

$$a^* = a\sqrt{-1+h^2/c^2}, \qquad b^* = b\sqrt{-1+h^2/c^2}.$$

De aquí se infiere que el plano secante $z = h$ empieza a cortar el hiperboloide de dos hojas sólo cuando $|h| \geqslant c$. En la capa entre los planos $z = -c$ y $z = +c$ no hay puntos de la superficie en con-

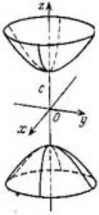

Fig. 97.3.

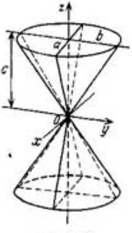

Fig. 97.4.

sideración. En virtud de la simetría respecto del plano Oxy, la superficie consta de dos hojas dispuestas fuera de la capa citada. Las secciones obtenidas como resultado del corte del hiperboloide por los planos Oyz y Oxz representan en sí unas hipérbolas.

I.6. *El número a_0 es igual a cero; los signos de λ_1, λ_2 coinciden y son contrarios al de λ_3.* Tenemos

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0. \tag{97.7}$$

La superficie definida por esta ecuación se llama *cono elíptico* (fig. 97.4) y la propia ecuación, ecuación *canónica* del cono elíptico. Los planos de coordenadas sirven de *planos de simetría* y el origen de coordenadas es *el centro de simetría*. Las líneas de intersección L_h del cono elíptico con los planos $z = h$ representan en sí elipses. Si el punto $M(x_0, y_0, z_0)$ se dispone en la superficie del cono, entonces las coordenadas del punto $M_t(tx_0, ty_0, tz_0)$, para todo número t, satisfacen la ecuación (97.7). Por consiguiente, toda la recta que pasa por el punto M_0 y el origen de coordenadas se halla *íntegramente* en la superficie dada.

Pasemos ahora a considerar la segunda ecuación de (95.14). Tenemos

$$\lambda_1 x^2 + \lambda_2 y^2 + b_0 z = 0.$$

II.7. *Los números λ_1, λ_2 son de un mismo signo.* Sin limitar la generalidad, podemos considerar que b_0 tiene signo opuesto, ya que al coincidir los signos de b_0 y λ_1, λ_2 obtenemos una superficie dispuesta simétricamente respecto del plano Oxy. El cambio habitual de los coeficientes nos da

$$z=\frac{x^2}{a^2}+\frac{y^2}{b^2}. \tag{97.8}$$

La superficie descrita por esta ecuación se llama *paraboloide elíptico* (fig. 97.5) y la propia ecuación, ecuación *canónica* del paraboloide elíptico. Para esta superficie Oxz y Oyz son *los planos de simetría*, el centro de simetría no existe. El paraboloide elíptico se

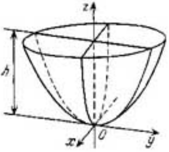

Fig. 97.5.

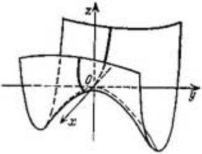

Fig. 97.6.

dispone en el semiespacio $z \geqslant 0$. Las líneas de intersección L_h del paraboloide elíptico con los planos $z = h$, $h > 0$, representan en sí unas elipses cuyas proyecciones sobre el plano Oxy se definen por la ecuación

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

donde $a^* = a\sqrt{h}$, $b^* = b\sqrt{h}$. De aquí se deduce que al crecer h, las elipses aumentan ilimitadamente, es decir, el paraboloide elíptico representa en sí una taza infinita.

Las secciones que se obtienen al cortar el paraboloide elíptico por los planos $y = h$ y $x = h$, representan en sí unas parábolas. Por ejemplo, el plano $x = h$ interseca la superficie a lo largo de la parábola

$$z-\frac{h^2}{a^2}=\frac{y^2}{b^2},$$

dispuesta en el plano $x = h$.

II.8. *Los números λ_1, λ_2 son de signos diferentes.* La superficie tipo para este caso se determina por la ecuación

$$z=\frac{x^2}{a^2}-\frac{y^2}{b^2}.$$

La superficie descrita por esta ecuación se llama *paraboloide hiperbólico* (fig. 97.6) y la propia ecuación, ecuación *canónica* del paraboloide hiperbólico. Los planos Oxz y Oyz son los planos de simetría, el centro de simetría no existe. Las líneas de intersección del paraboloide hiperbólico con los planos $z = h$ representan en sí, para $h > 0$, las hipérbolas

$$\frac{x^2}{a^{*2}} - \frac{y^2}{b^{*2}} = 1,$$

donde $a^* = a\sqrt{h}$, $b^* = b\sqrt{h}$, y, para $h < 0$, las hipérbolas

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

donde $a^* = a\sqrt{-h}$, $b^* = b\sqrt{-h}$. El plano $z = 0$ corta el paraboloide hiperbólico a lo largo de dos rectas

$$y = \pm \frac{b}{a} x.$$

Todas las superficies definidas por las ecuaciones **III—V** de (95.14) no dependen de z. Por ello, las proyecciones de las líneas de intersección de dichas superficies con los planos $z = h$ sobre el plano Oxy tampoco dependen de h. Las superficies de tal género se llaman *cilindros*, añadiéndose la definición *elíptico*, *hiperbólico*, etc., según sea la forma de la proyección de la superficie sobre el plano Oxy.

TEOREMA 97.1. *Por todo punto del hiperboloide de una hoja y del paraboloide hiperbólico pasan dos rectas diferentes, dispuestas íntegramente en las superficies indicadas.*

DEMOSTRACION. Examinemos un hiperboloide de una hoja definido por su ecuación canónica

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \tag{97.10}$$

Con cualesquiera α, β distintos de cero a la vez, un par de planos

$$\alpha\left(\frac{x}{a} + \frac{z}{c}\right) = \beta\left(1 - \frac{y}{b}\right), \qquad \beta\left(\frac{x}{a} - \frac{z}{c}\right) = \alpha\left(1 + \frac{y}{b}\right) \tag{97.11}$$

determina cierta recta Γ. Es fácil comprobar que la recta dada Γ se dispone íntegramente en la superficie (97.10). Más aún, por todo punto de esta superficie pasa una recta perteneciente a la familia de Γ.

En efecto, consideraremos (97.11) como un sistema de dos ecuaciones

$$\alpha\left(\frac{x}{a} + \frac{z}{c}\right) - \beta\left(1 - \frac{y}{b}\right) = 0,$$
$$\alpha\left(1 + \frac{y}{b}\right) - \beta\left(\frac{x}{a} - \frac{z}{c}\right) = 0$$

respecto de α, β. El determinante del sistema es igual a cero cuando, y sólo cuando, el punto $M(x, y, z)$ se dispone en el hiperboloide (97.10). Además, el rango de la matriz del sistema es igual a uno a ciencia cierta. Por consiguiente, α y β se determinan, salvo la proporcionalidad. Pero, esto significa precisamente la unicidad de la recta Γ que pasa por todo punto del hiperboloide.

Del modo análogo nos convencemos de que por todo punto del hiperboloide pasa una única resta Γ^*, definida por los planos

$$\nu\left(\frac{x}{a}+\frac{z}{c}\right)=\lambda\left(1+\frac{y}{b}\right), \qquad \lambda\left(\frac{x}{a}-\frac{z}{c}\right)=\nu\left(1-\frac{y}{b}\right).$$

Las rectas Γ y Γ^* son distintas. Los mismos razonamientos muestran que un hiperboloide hiperbólico

$$z=\frac{x^2}{a^2}-\frac{y^2}{b^2}$$

está cubierto por dos familias diferentes de rectas Π y Π^* las cuales vienen definidas por los planos

$$\alpha z=\beta\left(\frac{x}{a}+\frac{y}{b}\right), \qquad \beta=\alpha\left(\frac{x}{a}-\frac{y}{b}\right)$$

y

$$\nu z=\lambda\left(\frac{x}{a}-\frac{y}{b}\right), \qquad \lambda=\nu\left(\frac{x}{a}+\frac{y}{b}\right).$$

Ejercicios.

1. ¿Qué representa en sí un hiperplano diametral conjugado de una dirección dada, para las superficies de segundo grado?
2. Escríbanse las ecuaciones de un plano tangente para diferentes superficies de segundo grado.
3. Investíguense las propiedades ópticas de las superficies de segundo grado.

CAPÍTULO 12 EXPACIOS BILINEALES MÉTRICOS

§ 98. Matriz y determinante de Gram

Supongamos que en un espacio lineal K_n, definido sobre el campo numérico P se ha introducido cierta forma bilineal $\varphi(x, y)$. El espacio K_n se llama *bilineal métrico*, si a cada par de vectores x, y de K_n se le ha puesto en correspondencia un número (x, y) de P denominado producto escalar, con la particularidad de que

$$(x, y) = \varphi(x, y).$$

Si una forma bilineal en el espacio complejo K_n es hermitiana, K_n se llama espacio *bilineal métrico hermitiano*. En estos casos diremos también que en el espacio lineal se ha introducido una métrica bilineal.

Puede observarse cierta analogía entre los espacios bilineales métricos y los espacios euclídeos y unitarios, considerados anteriormente. No obstante, indiquemos ahora mismo algunas diferencias sustanciales. Al comparar las definiciones del producto escalar en los espacios euclídeo y unitario con la definición de la forma bilineal, no es difícil advertir que en los espacios bilineales métricos el mencionado producto escalar puede no ser, en el caso general, simétrico y definido positivo.

El estudio de los espacios euclídeos y unitarios se reducía a la investigación de las propiedades adicionales tanto de los propios espacios como de los operadores que actúan en los espacios y surgen con relación a las formas bilineales las cuales determinan los productos escalares. El problema de estudio de los espacios bilineales métricos es el mismo. La necesidad de introducir una definición debilitada del producto escalar es debida al hecho de que no siempre las funciones bilineales, estudiadas en conjunto con los vectores del espacio y los operadores, poseen la propiedad de simetría y de definición positiva.

Muchas definiciones y hechos serán iguales tanto para los espacios bilineales métricos ordinarios como para los bilineales hermitianos. Por esta razón, siempre cuando no haya lugar para equivocaciones, omitiremos en lo sucesivo la palabra «hermitiano» y realizaremos los cálculos correspondientes sólo para los espacios bilineales, sobreentendiendo tácitamente que para los espacios hermitianos los cálculos se efectúan de manera análoga.

El procedimiento principal para investigar cualesquiera espacios lineales consiste en la descomposición de un vector según el sistema dado de vectores y en el estudio de dicha descomposición en función de diferentes factores. En un espacio lineal general no existe un instrumento, con ayuda del cual se podría encontrar la descomposición, pero resulta que en el espacio bilineal métrico como tal instrumento actúa el producto escalar.

Tomemos un sistema arbitrario de vectores $x_1, x_2, \ldots, x_m$ del espacio bilineal métrico K_n y un vector $x \in K_n$. Veamos qué nos da la presencia del producto escalar para investigar la posibilidad de descomponer el vector x

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m \tag{98.1}$$

según un sistema dado. Al multiplicar escalarmente la igualdad (98.1), de manera sucesiva a la derecha, por $x_1, x_2, \ldots, x_m$, obtendremos, para determinar los coeficientes desconocidos $\alpha_1, \alpha_2, \ldots, \alpha_m$ de la descomposición (98.1), un sistema de ecuaciones algebraicas lineales

$$\begin{gathered}
\alpha_1(x_1, x_1) + \alpha_2(x_2, x_1) + \ldots + \alpha_m(x_m, x_1) = (x, x_1),\\
\alpha_1(x_1, x_2) + \alpha_2(x_2, x_2) + \ldots + \alpha_m(x_m, x_2) = (x, x_2),\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\alpha_1(x_1, x_m) + \alpha_2(x_2, x_m) + \ldots + \alpha_m(x_m, x_m) = (x, x_m).
\end{gathered} \tag{98.2}$$

La matriz G, que es una matriz transpuesta de este sistema, tiene la forma

$$G = \begin{pmatrix} (x_1, x_1) & (x_1, x_2) & \ldots & (x_1, x_m) \\ (x_2, x_1) & (x_2, x_2) & \ldots & (x_2, x_m) \\ \ldots & \ldots & \ldots & \ldots \\ (x_m, x_1) & (x_m, x_2) & \ldots & (x_m, x_m) \end{pmatrix} \tag{98.3}$$

y se denomina *matriz de Gram* del sistema de vectores $x_1, x_2, \ldots, x_m$. *El determinante* de esta matriz $G(x_1, x_2, \ldots, x_m)$ lleva el nombre *de Gram*. De este modo, el problema de investigación de las descomposiciones (98.1) está estrechamente vinculado con el problema de resolución de los sistemas (98.2).

Si los vectores $x_1, x_2, \ldots, x_m$ forman una base del espacio, la matriz de Gram será, para ellos, matriz de la forma bilineal principal (x, y). Las matrices de Gram para diferentes bases son congruentes y, por tanto, tienen rangos iguales. El rango de las matrices de Gram es un invariante del espacio bilineal métrico y se llama *rango* de dicho espacio. La diferencia entre la dimensión y el rango del espacio se denomina *defecto* del espacio. Si el defecto es distinto de cero, el espacio bilineal métrico se llamará *degenerado*. La regularidad de la forma bilineal principal implica la regularidad de las matrices de Gram para todas las bases. En este caso el espacio bilineal métrico se denominará *regular*. Para un espacio regular el sistema (98.2), donde $x_1, \ldots, x_m$ es la base, tiene siempre una solución única, puesto que la matriz del sistema será regular y se obtiene la posibilidad de investigar los coeficientes de la descomposición (98.1) como solución del sistema (98.2).

Supongamos que para ciertos vectores x, y de un espacio bilineal métrico K_n se verifica la igualdad $(x, y) = 0$. En este caso el vector y se llama *ortogonal a la derecha* respecto del vector x, y el vector x, *ortogonal a la izquierda* respecto del vector y. En los espacios bilineales métricos hemos de distinguir las ortogonalidades a la izquierda y a la derecha, puesto que en el caso general $(x, y) \neq (y, x)$. Si, en cambio, $(x, y) = (y, x) = 0$, los vectores se llamarán simplemente *ortogonales*. Teniendo en cuenta la linealidad del producto escalar respecto de cada argumento, es fácil comprobar que la ortogonalidad del vector y a la derecha respecto de los vectores $x_1, x_2, \ldots, x_m$ origina su ortogonalidad a la derecha respecto de su cualquier combinación lineal. Lo mismo puede decirse respecto de la ortogonalidad a la izquierda. De aquí proviene, en particular, que para que un vector del espacio K_n sea ortogonal a todos los vectores del subespacio lineal L, es necesario y suficiente que sea ortogonal a los vectores de una base cualquiera del subespacio L.

LEMA 98.1. *Si la matriz de Gram del sistema de vectores $x_1, x_2, \ldots, x_m$ es degenerada, entonces existen tales vectores u, v (que son combinaciones lineales no triviales de los vectores $x_1, x_2, \ldots, x_m$) que u es ortogonal a la derecha y v, a la izquierda respecto de todos los vectores de la cápsula lineal de los vectores $x_1, x_2, \ldots, x_m$.*

DEMOSTRACIÓN. Si la matriz de Gram es degenerada, sus filas son linealmente dependientes. Por consiguiente, existen tales números $\gamma_1, \gamma_2, \ldots, \gamma_m$, no todos iguales a cero, que la combinación lineal de las filas será nula, es decir,

$$\begin{aligned}
\gamma_1 (x_1, x_1) + \gamma_2 (x_2, x_1) + \ldots + \gamma_m (x_m, x_1) &= 0, \\
\gamma_1 (x_1, x_2) + \gamma_2 (x_2, x_2) + \ldots + \gamma_m (x_m, x_2) &= 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\gamma_1 (x_1, x_m) + \gamma_2 (x_2, x_m) + \ldots + \gamma_m (x_m, x_m) &= 0.
\end{aligned} \tag{98.4}$$

Si designamos

$$v = \sum_{j=1}^{n} \gamma_j x_j,$$

las correlaciones (98.4) significan que $(v, x_j) = 0$ para todo j. El vector v es una combinación lineal no trivial de los vectores x_1, $x_2, \ldots, x_m$, el mismo es ortogonal a la izquierda respecto a cualquiera de los vectores citados, por lo cual es ortogonal a todo vector de su cápsula lineal. El vector u se construye análogamente, pero siempre partiendo de la dependencia lineal de las columnas de la matriz de Gram.

COROLARIO. *Si la matriz de Gram para un sistema de vectores linealmente independiente es degenerada, la forma cuadrática* (x, x) *tiene un vector isótropo que pertenece a la cápsula lineal del sistema dado y es ortogonal a la derecha (a la izquierda) respecto de todos los vectores de esta cápsula.*

Efectivamente, en virtud de que los vectores del sistema son linealmente independientes, los vectores u, v serán no nulos; además, $(u, u) = (v, v) = 0$.

En varios casos de importancia el determinante de Gram sirve de medio muy cómodo para establecer el hecho de dependencia lineal o independencia lineal de un sistema de vectores.

LEMA 98.2. *Para todo sistema de vectores linealmente independiente el determinante de Gram es igual a cero.*

DEMOSTRACION. Sea un sistema $x_1, x_2, \ldots, x_m$ linealmente dependiente En este caso podemos representar el vector nulo x en forma de una combinación lineal no trivial de los vectores $x_1, x_2, \ldots$ $\ldots, x_m$. Pero, entonces, el sistema homogéneo (98.2) debe tener una solución no nula. Por consiguiente, el determinante de la matriz de este sistema, es decir, el determinante de Gram del sistema $x_1, x_2, \ldots, x_m$ será igual a cero.

TEOREMA 98.1. *Si una forma cuadrática* (x, x) *no tiene vectores isótropos, el determinante de Gram no es nulo, cuando, y sólo, cuando, su sistema de vectores es linealmente independiente.*

DEMOSTRACION. NECESIDAD. Supongamos que el determinante de Gram del sistema de vectores $x_1, x_2, \ldots, x_m$ no es igual a cero. Si suponemos que este sistema es linealmente dependiente, entonces, de acuerdo con el lema 98.2, el determinante de Gram debe ser nulo, lo que es imposible por hipótesis.

SUFICIENCIA. Supongamos que el sistema de vectores es linealmente independiente. Si el determinante de Gram es nulo, entonces, conforme al corolario del lema 98.1, debe existir un vector isótropo. Pesto que esto último es imposible según la hipótesis, el determinante de Gram no es igual a cero.

COROLARIO. *Si una forma cuadrática* (x, x) *es estrictamente de signo constante, entonces el determinante de Gram es igual a cero cuando, y sólo cuando, el sistema de vectores es linealmente dependiente.*

COROLARIO. *Si una forma bilineal* (x, x) *es simétrica y la forma cuadrática* (x, x) *es estrictamente de signo constante, entonces para cualesquiera dos vectores* x, y *se verifica la desigualdad de Cauchy—Bunia-*

kovski

$$|(x, y)|^2 \leqslant (x, x)(y, y), \tag{98.5}$$

con la particularidad de que la igualdad se alcanza cuando, y sólo cuando, los vectores x, y son linealmente dependientes.

A condiciones de esta afirmación el determinante de Gram para los vectores linealmente independientes x, y será positivo, de conformidad con el criterio de Sylvester o corolario del mismo, e igual a cero, para los vectores linealmente dependientes, de acuerdo con el lema 98.2. En ambos casos la desigualdad (98.5) tiene lugar. Si, en cambio, en (98.5) se alcanza una igualdad, los vectores x, y serán linealmente dependientes, en concordancia con el corolario anterior, puesto que será nulo su determinante de Gram.

Consideraremos algunas propiedades del determinante de Gram que son sencillas, pero bastante importantes. Estas propiedades no sólo generan numerosos corolarios, sino permiten frecuentemente atribuirles una clara interpretación geométrica.

PROPIEDAD 1. *El determinante de Gram no varía, al cambiar de lugar cualesquiera dos vectores en el sistema $x_1, x_2, \ldots, x_m$.*

En efecto, si en el sistema $x_1, x_2, \ldots, x_m$ cambiamos de lugar cualesquiera dos vectores x_i y x_j, en el determinante de Gram cambiarán de lugar entre sí la i-ésima y la j-ésima columnas, como también la i-ésima y la j-ésima filas. Además, el determinante de Gram cambiará de signo dos veces y, como resultado, quedará inalterable.

PROPIEDAD 2. *El determinante de Gram no varía cuando a un vector cualquiera del sistema $x_1, x_2, \ldots, x_m$ se le adiciona cualquier combinación lineal de los demás vectores.*

Evidentemente, es suficiente considerar un caso en que varía el vector x_1, puesto que todos los casos restantes se reducen, teniendo presente la propiedad 1, a este primer caso. Supongamos que al vector x_1 se le agrega el vector $\alpha_2 x_2 + \ldots + \alpha_m x_m$. Supongamos además, que la forma bilineal (x, y) es ordinaria. Es fácil comprobar, que el nuevo determinante de Gram se obtiene del inicial, sumando a la primera fila la fila segunda multiplicada por α_2, etc. hasta la última fila multiplicada por α_m, y a la primera columna la columna segunda multiplicada por α_2, etc., hasta la última columna multiplicada por α_m. Como resultado de tal procedimiento, según se sabe, el determinante no varía. Si la forma bilineal (x, y) es hermitiana, las columnas se multiplican por $\bar{\alpha}_2, \ldots, \bar{\alpha}_m$.

PROPIEDAD 3. *Si un vector del sistema $x_1, x_2, \ldots, x_m$ se multiplica por el número α, entonces el determinante de Gram queda multiplicado por α^2, siempre que la forma bilineal (x, y) sea ordinaria, y por $|\alpha|^2$, si la forma (x, y) sea hermitiana.*

Esta vez también resulta suficiente considerar el caso de variación del vector x_1. Pero la multiplicación del vector x_1 por el número α conduce a la multiplicación de la primera fila y la primera columna del determinante de Gram por el número α sólo en el caso en que la forma bilineal (x, y) sea ordinaria. En cambio, si la forma (x, y) es hermitiana, entonces la primera fila del determinante de Gram queda multiplicada por el número α, mientras que la primera columna, por el número $\bar{\alpha}$. De aquí precisamente proviene la propiedad enunciada.

PROPIEDAD 4. *Si cada uno de los vectores $x_1, x_2, \ldots, x_m$ es ortogonal a la izquierda (a la derecha) respecto de todos los vectores que le anteceden, entonces para el determinante de Gram se verifica la igualdad*

$$G(x_1, x_2, \ldots, x_m) = \prod_{i=1}^{m} (x_i, x_i). \tag{98.6}$$

Efectivamente, la ortogonalidad a la izquierda (a la derecha) de cada uno de los vectores del sistema $x_1, x_2, \ldots, x_m$ respecto de todos los vectores antecedentes lleva a que la matriz de Gram será triangular derecha (izquierda). Pero el determinante de una matriz triangular es igual al producto de los elementos diagonales, de donde se infiere (98.6).

Son de mayor interés las propiedades de la matriz y del determinante de Gram en aquellos casos cuando la forma bilineal (x, y) interviene como simétrica real o simétrica hermitiana y es definida positiva. Desde luego, estos casos significan nada más que el espacio bilineal métrico K_n es, de hecho, euclídiano o bien, correspondientemente, unitario.

En un espacio euclídeo y unitario la matriz de Gram será, para cualquier sistema básico, una matriz definida positiva de forma cuadrática (x, x). De acuerdo con el criterio de Sylvester, todos los menores principales de la matriz de Gram serán positivos. Puesto que todo sistema de vectores linealmente independiente puede ser construido de modo que se obtenga una base, de aquí se deduce que será válido el

LEMA 98.3. *En un espacio euclídeo y unitario el determinante de Gram para cualquier sistema linealmente independiente de vectores es positivo.*

En un espacio euclídeo el determinante de Gram tiene una interpretación geométrica muy simple. De esto nos dice el

TEOREMA 98.2. *En un espacio euclídeo el determinante de Gram $G(x_1, x_2, \ldots, x_m)$ del sistema de vectores $x_1, x_2, \ldots, x_m$ es igual al cuadrado del volumen $V^2(x_1, x_2, \ldots, x_m)$ de dicho sistema de vectores.*

DEMOSTRACIÓN. Examinemos una función real $G^{1/2}(x_1, \ldots, x_m)$ de m argumentos vectoriales $x_1, x_2, \ldots, x_m$. La función satisface las propiedades A, B de (36.3), conforme a las propiedades 2, 3 del determinante de Gram. En un espacio euclídeo cada vector de cualquier sistema ortonormalizado de vectores es ortogonal a todos los vectores antecedentes del sistema. Por esto, en concordancia con (98.6), la función $G^{1/2}(x_1, x_2, \ldots, x_m)$ satisface también la condición C de (36.3). Pero ahora del teorema 36.1 se desprende que esta función coincide con el volumen del sistema de vectores.

COROLARIO. *Para todo sistema de vectores $x_1, x_2, \ldots, x_m$ de un espacio euclídeo se verifican las desigualdades*

$$0 \leqslant G(x_1, x_2, \ldots, x_m) \leqslant \prod_{i=1}^{m} (x_i, x_i),$$

con la particularidad de que la igualdad a la izquierda se alcanza cuando, y sólo cuando, el sistema de vectores es linealmente dependiente, mientras que la igualdad a la derecha, cuando, y sólo cuando, bien el sistema de vectores es ortogonal bien contiene el vector nulo.

La validez de esta afirmación proviene del primer corolario del teorema 98.1 y de la propiedad del volumen del sistema de vectores descrita por la desigualdad de Hadamard (36.1).

COROLARIO. *Para cualquier sistema de vectores $x_1, x_2, \ldots, x_m$ de un espacio euclídeo se verifica la desigualdad*

$$G(x_1, \ldots, x_l, x_{l+1}, \ldots, x_m) \leqslant$$
$$\leqslant G(x_1, \ldots, x_l) \cdot G(x_{l+1}, \ldots, x_m)$$

con la particularidad de que la igualdad tiene lugar cuando, y sólo cuando, o bien los conjuntos de vectores $x_1, \ldots, x_l$ y $x_{l+1}, \ldots, x_m$ son ortogonales o bien uno de los conjuntos citados representa un sistema linealmente dependiente.

La demostración se basa en un análisis muy sencillo de la fórmula (35.4). Recordemos sólo lo siguiente. Si $L_1 \subseteq L_2$, donde L_1, L_2 son unos subespacios cualesquiera, entonces $|\text{ort}_{L_2} x| \leqslant |\text{ort}_{L_1} x|$ para todo vector x. Con ello, la igualdad tiene lugar sólo en el caso cuando $x \perp L_2$.

Ejercicios.

1. ¿Serán equivalentes los problemas de la búsqueda de las descomposiciones (98.1) y la solución de los sistemas (98.2)?
2. ¿Qué da la solución del sistema (98.2), si el vector x no pertenece a la cápsula lineal de los vectores $x_1, \ldots, x_m$?
3. ¿Cómo se representa la matriz de Gram (98.3), si:
los vectores $x_1, \ldots, x_m$ son ortogonales dos a dos,
cada uno de los vectores $x_1, \ldots, x_m$ es ortogonal a la izquierda (a la derecha) respecto de todos los vectores antecedentes (posteriores),
cada uno de los vectores $x_1, \ldots, x_m$ es ortogonal a la izquierda (a la derecha) respecto de todos los vectores posteriores (antecedentes),
cada uno de los vectores $x_{l+1}, \ldots, x_m$ es ortogonal a la izquierda (a la derecha) respecto de cada uno de los vectores $x_1, \ldots, x_l$?
4. ¿Cómo varía la matriz de Gram cuando el sistema de vectores se somete a las transformaciones elementales?

5. Demuéstrese que si en un espacio bilineal métrico ordinario de la condición $(x, y) = 0$ se sigue siempre que $(y, x) = 0$, entonces el producto escalar está dado por cualquier forma bilineal, simétrica o antisimétrica.

6. ¿Será cierta la afirmación 5 para un espacio bilineal métrico hermitiano?

7. Sea G una matriz de Gram para cierta base en un espacio bilineal métrico K_n, regular hermitiano. Demuéstrese que para el operador U con la matriz $\overline{G}^{-1}G'$ en la misma base se verifica la igualdad

$$(Ux, Ux) = (x, x),$$

cualesquiera que sean los vectores $x \in K_n$.

8. Demuéstrese que para cualquier operador lineal A que actúa en el espacio euclídeo o unitario K_n la correlación

$$k(A) = \frac{G(Ax_1, \ldots, Ax_m)}{G(x_1, \ldots, x_m)}$$

no depende de los vectores $x_1, \ldots, x_m$ y es igual al producto de los cuadrados de módulos de los valores propios del operador A.

9. Demuéstrese que para todo sistema linealmente independiente de vectores $x_1, \ldots, x_m$ de un espacio euclídeo o unitario y todo vector z se verifica la desigualdad

$$\frac{G(x_1, \ldots, x_m, z)}{G(x_1, \ldots, x_m)} \leqslant \frac{G(x_1, \ldots, x_{m-1}, z)}{G(x_1, \ldots, x_{m-1})}.$$

§ 99. Subespacios regulares

Todo subespacio lineal L de K_n puede considerarse como espacio bilineal métrico respecto del mismo producto escalar que se ha introducido en K_n. En el caso general, de la regularidad de K_n no proviene la regularidad de L, y viceversa.

TEOREMA 99.1. *Para que en el espacio K_n sean regulares todos sus subespacios, es necesario y suficiente que la forma cuadrática (x, x) no tenga vectores isótropos.*

DEMOSTRACION. NECESIDAD. Supongamos que en N_n todos los subespacios son regulares. Entonces, serán regulares también todos los subespacios unidimensionales. Pero las matrices de Gram para los vectores no nulos x coinciden con el producto escalar (x, x), el cual debe ser distinto de cero, puesto que los subespacios unidimensionales son regulares.

SUFICIENCIA. Supongamos que $(x, x) \neq 0$ para todo $x \neq 0$. Consideremos un subespacio cualquiera L y una base en él $x_1, x_2, \ldots, x_m$. De conformidad con el teorema 98.1, el determinante de Gram para este sistema es distinto de cero, es decir, el subespacio L es regular.

COROLARIO. *Para que en un espacio bilineal métrico K_n sean regulares todos sus subespacios, es necesario y suficiente que todos sus subespacios unidimensionales sean regulares.*

COROLARIO. *En cualquier espacio bilineal métrico ordinario complejo existen subespacios degenerados unidimensionales.*

Para demostrar esta afirmación es suficiente recordar que en un espacio bilineal métrico ordinario complejo toda forma cuadrática tiene vectores isótropos.

Cuando una forma cuadrática tiene vectores isótropos, en el espacio bilineal métrico existirán tanto subespacios degenerados como regulares. Si la forma bilineal (x, y) es de rango r, está claro que no puede haber subespacios regulares cuya dimensión sea superior a r. Pero, los subespacios regulares de dimensión r existen. Como ejemplo podemos indicar el subespacio tendido sobre aquellos vectores de la base canónica, para los cuales la matriz de Gram coincide con la matriz M de (92.5).

Diremos que el conjunto de vectores F de un espacio bilineal métrico K_n *es ortogonal a la derecha, a la izquierda* o simplemente *ortogonal* al conjunto de vectores G de K_n, si para cada par de vectores x, y, donde $x \in F$, $y \in G$, se cumple una relación análoga de ortogonalidad. Está claro que la totalidad de todos los vectores del espacio K_n, ortogonales a la derecha (a la izquierda) respecto de cada uno de los vectores del conjunto F, es un subespacio. Se denomina *complemento ortogonal a la derecha (a la izquierda)* del conjunto F y se designa con $F^{\perp}$ $({}^{\perp}F)$.

En los espacios euclídeo y unitario los subespacios ${}^{\perp}K_n$ y $K_n^{\perp}$ coinciden y se componen sólo de un vector nulo. En los espacios bilineales métricos estos subespacios pueden ser diferentes y no es obligatorio que se compongan sólo de un vector nulo. Los subespacios ${}^{\perp}K_n$ y $K_n^{\perp}$ se denominan subespacios *nulos* en K_n, izquierdo y derecho, respectivamente.

Cabe notar que para todo conjunto de vectores F son siempre justas las inclusiones $K_n^{\perp} \subseteq F^{\perp}$, ${}^{\perp}K_n \subseteq {}^{\perp}F$, y para cualesquiera vectores de ${}^{\perp}K_n$ o $K_n^{\perp}$ las matrices de Gram resultan nulas.

TEOREMA 98.2. *Las dimensiones de los subespacios nulos izquierdo y derecho coinciden y son iguales al defecto de la forma bilineal* (x, y).

DEMOSTRACIÓN. Elijamos en K_n una base $x_1, x_2, \ldots, x_n$. Tomemos un vector arbitrario x de $K_n^{\perp}$ y representémoslo en forma de una descomposición según la base, por analogía con (98.1). La condición de pertenencia del vector x al subespacio $K_n^{\perp}$ es equivalente a las condiciones de ortogonalidad a la derecha del vector x respecto de cada uno de los vectores de la base. Pero estas condiciones conducen a la resolución del sistema homogéneo del tipo (98.2) con el fin de hallar los coeficientes de la descomposición. Se sabe (véase § 48) que el conjunto de soluciones de dicho sistema es un subconjunto cuya dimensión es igual al defecto de la matriz de Gram o, que es lo mismo, al defecto de la forma bilineal (x, y). La demostración para el subespacio nulo izquierdo se realiza de manera análoga.

COROLARIO. *Para que el espacio K_n sea regular, es necesario y suficiente que los subespacios derecho e izquierdo se compongan sólo del vector nulo.*

En los espacios euclídeo y unitario todo subespacio es ortogonal a su complemento ortogonal y determina la descomposición de todo el espacio no sólo en una recta, sino incluso en la suma ortogonal de estos subespacios. En los espacios bilineales métricos no siempre tienen lugar los hechos análogos.

TEOREMA 99.3. *Sea L un subespacio en K_n. Para que existan las descomposiciones*

$$K_n = L \dot{+} L^{\perp} = L \dot{+} {}^{\perp}L, \tag{99.1}$$

es necesario y suficiente que el espacio L sea regular.

DEMOSTRACION. NECESIDAD. Supongamos que las descomposiciones (99.1) tienen lugar. Consideraremos L como un espacio bilineal métrico con el mismo producto escalar que figuraba en K_n. La intersección $L \cap L^{\perp}$ es el subespacio nulo derecho en L. Puesto que las sumas (99.1) son directas, este subespacio sólo contiene el vector nulo. De acuerdo con el corolario del teorema 99.2, esto significa que el subespacio L es regular.

SUFICIENCIA. Si el subespacio L es regular, entonces la intersección $L \cap L^{\perp}$ contendrá sólo el vector nulo y resta por señalar que todo vector $x \in K_n$ puede ser representado en la forma $x = u + v$, donde $u \in L$, $v \in L^{\perp}$. Elijamos en L una base $x_1, x_2, \ldots, x_m$. Para que exista la descomposición buscada $x = u + v$, es necesario y suficiente que en L se encuentre tal vector u que $x - u$ sea ortogonal a la derecha respecto de los vectores $x_1, x_2, \ldots, x_m$. Esta vez también obtenemos un sistema de ecuaciones algebraicas lineales con la matriz de Gram para determinar los coeficientes de la descomposición del vector u según los vectores $x_1, x_2, \ldots, x_m$. La matriz citada es regular y el sistema tiene solución, es decir, el vector u existe.

Desde luego, todo lo que hemos dicho respecto al subespacio $L^{\perp}$ es válido por completo para el subespacio ${}^{\perp}L$.

COROLARIO. *Si un subespacio regular L tiene dimensión m, entonces la dimensión de los subespacios $L^{\perp}$ y ${}^{\perp}L$ es igual a $n-m$.*

Con miras a demostrar esta afirmación resulta suficiente hacer uso de la igualdad (19.1) y recordar que la dimensión de los subespacios $L \cap L^{\perp}$ y $L \cap {}^{\perp}L$ es igual a cero.

COROLARIO. *Si un subespacio regular L tiene dimensión máxima, entonces $L^{\perp} = K_n^{\perp}$, ${}^{\perp}L = {}^{\perp}K_n$.*

En efecto, sea r el rango de la forma bilineal (x, y). Como ya se ha observado, el subespacio L será de dimensión r, mientras que los subespacios $K_n^{\perp}$ y ${}^{\perp}K$ tendrán la dimensión $n-r$. Pero los subespacios $K_n^{\perp}$ y ${}^{\perp}K_n$ tienen también esta misma dimensión $n - r$, y, además, $K_n^{\perp} \subseteq L^{\perp}$, ${}^{\perp}K_n \subseteq {}^{\perp}L$. Por eso, $K_n^{\perp} = L^{\perp}$, ${}^{\perp}K_n = {}^{\perp}L$.

En cuanto a las descomposiciones del tipo (99.1) en sumas ortogonales, cabe notar que del teorema 99.3 proviene el

COROLARIO. *Sea L un subespacio regular de dimensión máxima. Las descomposiciones* (99.1) *serán ortogonales cuando, y sólo cuando, los subespacios nulos izquierdo y derecho coinciden.*

Efectivamente, si las descomposiciones (99.1) son ortogonales, entonces $L^{\perp}$ es ortogonal a L no sólo a la derecha sino también a la izquierda, es decir, $L^{\perp} \subseteq {}^{\perp}L$. Por analogía, tenemos ${}^{\perp}L \subseteq L^{\perp}$. Por consiguiente, $L^{\perp} = L^{\perp}$. El hecho de que L es de dimensión máxima significa que $K_n^{\perp} = {}^{\perp}K_n$. En cambio, si los subespacios nulos coinciden, de aquí se infiere que $L^{\perp} = {}^{\perp}L$, es decir, los subespacios $L^{\perp}$ y ${}^{\perp}L$ son ortogonales a L tanto a la derecha como a la izquierda y las descomposiciones (98.5) son ortogonales.

Ahora podemos dar la respuesta a la pregunta sobre la relación existente entre la descomposición (98.1) y la solución del sistema (98.2). Supongamos que los vectores $x_1, \ldots, x_m$ forman la base de un subespacio regular L. De acuerdo con el teorema 99.3, tienen lugar las descomposiciones directas (99.1). Por ello, todo vector x del espacio bilineal métrico K_n puede ser representado de manera única en la forma $x = u + v$, donde $u \in L$, $v \in {}^{\perp}L$. Recordemos que el vector u se llama proyección del vector x sobre el subespacio L paralelamente al subespacio ${}^{\perp}L$. Si resolvemos el sistema de ecuaciones algebraicas lineales (98.2) y componemos el vector

$$u = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m, \tag{99.2}$$

precisamente este vector será la proyección de x sobre el subespacio L paralelamente al subespacio ${}^{\perp}L$. Efectivamente, el vector u pertenece a L, mientras que la diferencia $x - u$, de conformidad con (98.2), es ortogonal a la izquierda respecto de los vectores $x_1, \ldots, x_m$. Por lo tanto, $x - u$ pertenece a ${}^{\perp}L$. Está claro que para proyectar el vector x sobre el subespacio L paralelamente a $L^{\perp}$, es necesario resolver el siguiente sistema

$$\begin{gathered}
\alpha_1 (x_1, x_1) + \alpha_2 (x_1, x_2) + \ldots + \alpha_m (x_1, x_m) = (x_1, x), \\
\alpha_1 (x_2, x_1) + \alpha_2 (x_2, x_2) + \ldots + \alpha_m (x_2, x_m) = (x_2, x), \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\alpha_1 (x_m, x_1) + \alpha_2 (x_m, x_2) + \ldots + \alpha_m (x_m, x_m) = (x_m, x)
\end{gathered} \tag{99.3}$$

y después calcular la proyección buscada según (99.2). En el caso de un espacio bilineal métrico hermitiano los coeficientes α_j en (92.2) se sustituyen por $\bar{\alpha}_j$.

Ejercicios.

1. Descríbanse todos los subespacios regulares de dimensión máxima.
2. Demuéstrese que para todo conjunto L tienen lugar las inclusiones

$$L \subseteq {}^{\blacktriangle}(L^{\blacktriangle}), \qquad L \subseteq ({}^{\perp}L)$$

¿En qué casos se alcanzan las igualdades en estas fórmulas?

3. Demuéstrese que si L es un subespacio regular de dimensión máxima en el espacio K_n, entonces

$$^{\perp}(L\perp) = (\perp L)\perp = K_n.$$

4. Demuéstrese que si un producto escalar está dado mediante una forma bilineal simétrica o antisimétrica, entonces para cualquier conjunto F se verifica la igualdad $F\perp = \perp F$.

5. ¿De qué modo están ligadas entre sí la descomposición (98.1) y la solución del sistema (98.2), si la matriz de Gram del sistema $x_1, x_2, \ldots, x_m$ es degenerada?

6. ¿Puede existir en un espacio regular una base compuesta por los vectores isótropos?

7. ¿Qué puede decirse sobre un producto escalar, si las proyecciones sobre el subespacio fijado L paralelamente a $\perp L$ y $L\perp$ coinciden para todos los vectores?

8. ¿Qué puede decirse sobre un producto escalar, si las proyecciones de un vector fijado sobre todos los subespacios L paralelamente a $\perp L$ y $L\perp$ coinciden?

9. Sea L un subespacio regular del espacio bilineal métrico hermitiano K_n de rango $r < n$. Demuéstrese la equivalencia de las siguientes afirmaciones:
el subespacio $\perp L$ es de dimensión $n < r$,
el subespacio $L\perp$ es de dimensión $n - r$,
el subespacio L es de dimensión r,
los subespacios $L\perp$ y $K_n^{\perp}$ coinciden,
los subespacios $\perp L$ y $\perp K_n$ coinciden,
el subespacio $\perp L$ se compone de los vectores isótropos y el vector nulo,
el subespacio $L\perp$ se compone de los vectores isótropos y el vector nulo,
el producto escalar en el subespacio $\perp L$ es igual a cero,
el producto escalar en el subespacio $L\perp$ es igual a cero.

10. ¿Qué forma tendrá la matriz de Gram para las bases compuestas por las bases de un subespacio regular L y un subespacio $\perp L$ $(L\perp)$?

§ 100. Ortogonalidad en las bases

En los espacios bilineales métricos las bases no son de igual paridad. Entre tales espacios hay algunos, para los cuales los sistemas (98.2) se resuelven y se utilizan con una facilidad singular. Por ejemplo, en el caso en que la parte considerable de la matriz de Gram se compone de los elementos nulos. Según el tipo que tengan las matrices de Gram, consideraremos varias clases de bases en los espacios bilineales métricos.

Las matrices más sencillas son diagonales. Las matrices diagonales de Gram aparecen cuando, y sólo cuando, las bases se componen de vectores ortogonales dos a dos. Tales bases se denominarán *ortogonales*. Un sistema de vectores que forman una base ortogonal en su cápsula lineal se llamará *sistema ortogonal*.

Las bases ortogonales pueden definirse de diferente modo. La definición mediante la ortogonalidad dos a dos no siempre es cómoda para la comprobación, sobre todo en los casos en que los vectores de la base se construyen sucesivamente, a partir del primero. Por eso, resulta a veces útil emplear la siguiente definición.

Una base $e_1, e_2, \ldots, e_n$ se llama ortogonal, si cada uno de sus vectores es ortogonal a todos los vectores antecedentes.

La matriz de Gram para los vectores que satisfacen esta definición es diagonal, por lo cual ambas definiciones son equivalentes. Generalmente, en la base pueden haber tanto vectores no isótropos, como isótropos. Los vectores de una base ortogonal siempre pueden conmutarse de modo tal que los vectores no isótropos vayan primeros y los isótropos sean últimos. La forma diagonal de la matriz de Gram en este caso queda, naturalmente, inalterable.

No todo espacio bilineal métrico o espacio bilineal métrico hermitiano tiene las bases ortogonales. Si existe aunque sea una sola base ortogonal, esto es testimonio de que la matriz de la forma bilineal (x, y) es diagonal en la base dada. Por consiguiente, la matriz de la forma bilineal (x, y) en cualquier otra base debe ser congruente de la diagonal. Por supuesto, la afirmación inversa es también cierta. Por esto

Para que en un espacio bilineal métrico o en un espacio bilineal métrico hermitiano exista una base ortogonal, es necesario y suficiente que la matriz de la forma bilineal (x, y) *sea congruente de la matriz diagonal. En este caso el conjunto de todas las bases ortogonales coincide, salvo la permutación de los vectores, con el conjunto de las bases canónicas de la forma bilineal* (x, y).

Ahora, apoyándonos en las investigaciones efectuadas anteriormente de las formas bilineales, podemos decir que entre los espacios bilineales métricos ordinarios tienen bases ortogonales aquellos espacios y sólo aquellos en los que la forma bilineal básica (x, y) es simétrica. Entre los espacios bilineales métricos hermitianos tienen bases ortogonales aquellos que cuentan con una forma bilineal básica (x, y) hermitiana o antihermitiana, como también con la forma bilineal (x, y) que tiene la parte real o imaginaria de signo constante de la forma cuadrática (x, x).

Indiquemos ahora mismo una distinción de principio que existe entre los espacios bilineales métricos con bases ortogonales, ordinarios y hermitianos. En un espacio bilineal métrico ordinario K_n la presencia de una base ortogonal lleva tras de sí la simetría del producto escalar (x, y) y esto último asegura, a su vez, la existencia de una base ortogonal en cualquier subespacio de K_n. En un espacio bilineal métrico hermitiano, del hecho de que en el mismo existe una base ortogonal automáticamente no se desprende, en el caso general, la existencia de una base ortogonal en cualquiera de sus subespacios. No obstante, si el producto escalar está dado mediante una forma bilineal simétrica hermitiana o antisimétrica hermitiana, este corolario sigue siendo válido.

Consideraremos una base ortogonal cualquiera $e_1, e_2, \ldots, e_n$ del espacio bilineal métrico K_n. En esta base hay tantos vectores isótropos y tantos no isótropos cuales son el defecto y el rango, respectivamente, del espacio K_n. Tomando en consideración la ley de inercia de las formas cuadráticas, concluimos que si la forma bilineal

(x, y) es simétrica real o simétrica hermítiana, entonces toda base ortogonal tendrá el mismo número de vectores con valores positivos y negativos de las magnitudes (e_i, e_i). Estos números son invariantes para todas las bases ortogonales en K_n. Con arreglo a esto, hablaremos del índice positivo y negativo, como también de la signatura de los espacios con la forma simétrica (x, y). En el caso de los espacios bilineales métricos con la forma asimétrica (x, y), se tratarán sólo el rango y el defecto de los espacios.

Si el espacio K_n es regular, cada base ortogonal $e_1, e_2, \ldots, e_n$ está privada de vectores isótropos. En este caso para todo vector $x \in K_n$ es válida la descomposición

$$x = \sum_{j=1}^{n} \frac{(x, e_j)}{(e_j, e_j)} e_j. \tag{100.1}$$

Efectivamente, al multiplicar sucesiva y escalarmente la igualdad

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n \tag{100.2}$$

a la derecha por los vectores $e_1, e_2, \ldots, e_n$, obtendremos

$$\alpha_j = \frac{(x, e_j)}{(e_j, e_j)}$$

para todo j. Los vectores de la base ortogonal en un espacio regular pueden normalizarse obteniéndose la base *ortonormalizada*. Para la base ortonormalizada $e_1, e_2, \ldots, e_n$ se cumplen las correlaciones $|(e_j, e_j)| = 1$, cualquiera que sea j.

En los espacios degenerados entre los vectores de cualquier base habrá necesariamente vectores isótropos. Por esta razón, la representación (100.1) para la descomposición (100.2) de los vectores del espacio ya no será válida. No obstante, en estos espacios también las bases ortogonales resultan ser bastante útiles. Como ejemplo de su empleo probemos que es lícito el

TEOREMA 100.1. *Si en un espacio provisto de un producto escalar existe una base ortogonal, los subespacios nulos derecho e izquierdo coinciden.*

DEMOSTRACIÓN. Supongamos que en el espacio K_n de rango r existe una base ortogonal $e_1, e_2, \ldots, e_n$. Convengamos en considerar que los vectores $e_1, \ldots, e_r$ son no isótropos, mientras que $e_{r+1}, \ldots, e_n$ son isótropos. Tomemos arbitrariamente un vector $x \in K_n$ y descompongámoslo de acuerdo con (100.2). Haciendo uso de la representación (100.2) y tomando en consideración la ortogonalidad de la base e isotropía de los vectores $e_{r+1}, \ldots, e_n$, es fácil establecer que $(x, e_j) = (e_j, x) = 0$ para $r < j \leqslant n$. Por consiguiente, los vectores $e_{r+1}, \ldots, e_n$ figuran simultáneamente tanto en el subespacio nulo derecho como en el nulo izquierdo. Pero los vectores $e_{r+1}, \ldots, e_n$ son linealmente independientes, como vectores de la base, y su número equivale a la dimensión de los subespacios nulos, por lo cual ambos subespacios nulos coinciden.

COROLARIO. *Si en un espacio bilineal métrico el producto escalar viene dado por una forma bilineal simétrica o simétrica hermitiana, los subespacios nulos derecho e izquierdo coinciden.*

COROLARIO. *En toda base ortogonal los vectores isótropos, y sólo ellos, forman una base del subespacio nulo común.*

COROLARIO. *Si en un espacio provisto de un producto escalar existe una base ortogonal, el espacio puede ser descompuesto en una suma ortogonal de cualquier subespacio regular de dimensión máxima y un subespacio nulo.*

El último corolario significa, de hecho, que el estudio de cualesquiera espacios degenerados con bases ortogonales se reduce al estudio por separado de los subespacios regulares con bases ortogonales y de los subespacios, donde el producto escalar es igual a cero.

El conocer la base ortogonal en un espacio permite no sólo indicar la base ortogonal en el subespacio regular de dimensión máxima, sino también obtener la descomposición explícita de la proyección ortogonal de cualquier vector sobre dicho subespacio según la base ortogonal de éste. En efecto, sea $e_1, e_2, \ldots, e_n$ una base ortogonal en K_n, sean $e_1, \ldots, e_r$ los vectores no isótropos y $e_{r+1}, \ldots, e_n$, isótropos. Designemos con L el subespacio tendido sobre los vectores $e_1, \ldots, e_r$. Está claro que L es regular, tiene la dimensión máxima, $L^\perp = {}^\perp L$ y, además,

$$K_n = L \oplus L^\perp.$$

Todo vector x de K_n puede ser representado unívocamente en forma de la suma $x = u + v$, donde $u \in L$, $v \in L^\perp$. Aquí, u se llama proyección ortogonal izquierda del vector x sobre el subespacio L, mientras que v, perpendicular izquierda a dicho subespacio. Escribamos para x la descomposición (100.2) según la base $e_1, e_2, \ldots, e_n$. La fórmula (100.1) ya no es válida. No obstante, conviene notar que los primeros r sumandos en (100.2) forman el vector u y los últimos $n - r$ sumandos, el vector v. Al multiplicar la igualdad (100.2) sucesiva y escalarmente a la derecha por $e_1, \ldots, e_r$, llegamos a que

$$u = \sum_{j=1}^{r} \frac{(x, e_j)}{(e_j, e_j)} e_j.$$

La proyección v del vector x sobre el subespacio nulo se determina de una manera más simple

$$v = x - \sum_{j=1}^{r} \frac{(x, e_j)}{(e_j, e_j)} e_j.$$

Lo único que no se puede hacer ahora es hallar la descomposición del vector v según los vectores $e_{r+1}, \ldots, e_n$, haciendo uso del producto escalar, pese a que la propia descomposición existe.

Como ya se ha dicho, las bases ortogonales existen no en todo espacio bilineal métrico o espacio bilineal métrico hermitiano. Esta circunstancia nos obliga a buscar otras clases de las bases que sean más cómodas desde el punto de vista del producto escalar prefijado en el espacio. La solución se dicta por la expresión canónica para la matriz de la forma bilineal.

La base $e_1, e_2, \ldots, e_n$ se denomina *seudoortogonal*, si cada uno de sus vectores es ortogonal a la izquierda respecto de todos los vectores antecedentes y cada uno de sus vectores isótropos es ortogonal a la izquierda respecto de todos los vectores de la base. Un sistema de vectores que forman la base seudoortogonal en su cápsula lineal se llamará *seudoortogonal*.

Hemos de notar que en la definición dada la ortogonalidad de los vectores a la izquierda respecto de todos los anteriores puede sustituirse por la ortogonalidad de los vectores a la derecha respecto de todos los vectores posteriores. Esto determina las mismas condiciones.

La matriz de Gram para los vectores de una base seudoortogonal es *trapezoidal* derecha. Si los vectores de la base se permutan de modo tal que los vectores no isótropos vayan primeros y los isótropos sean últimos, entonces la matriz de Gram no sólo queda trapezoidal derecha, sino adquiere, además, la expresión canónica (92.5). Las investigaciones realizadas anteriormente, referentes a la reducción de la matriz de una forma bilineal a la expresión canónica, nos proporcionan la respuesta completa a la pregunta sobre las condiciones de existencia de una base seudoortogonal.

La base seudoortogonal existe en cualquier espacio bilineal métrico hermitiano, como también en todo espacio bilineal métrico ordinario, a excepción de los espacios con la forma bilineal antisimétrica (x, y). *El conjunto de todas las bases seudoortogonales coincide, salvo la permutación de los vectores, con el conjunto de las bases canónicas de la forma bilineal* (x, y).

Toda base ortogonal es seudoortogonal. En un espacio bilineal métrico ordinario no pueden existir a la vez una base ortogonal y una seudoortogonal que no sea ortogonal. Esto se debe a que la existencia de aunque sea una sola base ortogonal lleva trás de sí la simetría de todas las matrices de Gram. La matriz trapezoidal derecha puede ser simétrica sólo en el caso, si es diagonal. En un espacio bilineal métrico hermitiano pueden existir simultáneamente una base ortogonal y una seudoortogonal que no sea ortogonal. Esto significa que la matriz compleja trapezoidal derecha puede ser congruente según Hermite de la matriz diagonal, lo que se confirma también por el ejemplo (92.8).

Si el espacio K_n es regular, toda base seudoortogonal no tiene vectores isótropos, puesto que la matriz trapezoidal derecha puede ser regular sólo en el caso cuando respresenta en sí una matriz triangular derecha con elementos diagonales no nulos. En un espacio regular para los coeficientes α_j de la descomposición (100.2) del vector x según los vectores de la base seudoortogonal $e_1, e_2, \ldots, e_n$ obtenemos un sistema de ecuaciones algebraicas lineales con una matriz triangular izquierda. En efecto, multiplicando sucesivamente la igualdad (100.2) a la derecha por $e_1, e_2, \ldots, e_n$, encontramos que

$$\begin{aligned} &\alpha_1(e_1, e_1) &&= (x, e_1) \\ &\alpha_1(e_1, e_2) + \alpha_2(e_2, e_2) &&= (x, e_2) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\alpha_1(e_1, e_n) + \alpha_2(e_2, e_n) + \ldots + \alpha_n(e_n, e_n) &&= (x, e_n). \end{aligned} \tag{100.3}$$

De aquí determinamos sucesivamente $\alpha_1, \alpha_2, \ldots, \alpha_n$. Por supuesto en el espacio regular se puede normalizar los vectores de la base seudoortogonal y de este modo obtener la base seudoortonormalizada, para la cual $|(e_j, e_j)| = 1$, cualquiera que sea j.

Cabe notar que el proceso de resolución del sistema (100.3) da mucho más que simplemente la descomposición del vector x según la base seudoortonormalizada $e_1, e_2, \ldots, e_n$. Podemos calcular de paso, sin esfuerzos adicionales, todos los vectores

$$u_k = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k.$$

Los vectores u_k forman una sucesión de proyecciones de un mismo vector x sobre los subespacios encajados uno dentro del otro

$$L_1 \subseteq L_2 \subseteq \ldots \subseteq L_k,$$

donde L_k es la cápsula lineal de los vectores $e_1, e_2, \ldots, e_k$. Si consideramos u_k como "aproximación" a la solución x, la ortogonalidad a la izquierda del "error" $v_k = x - x_k$ respecto del subespacio L_k significa en realidad la ortogonalidad de v_k a la izquierda respecto de $u_1, u_2, \ldots, u_k$. Todas estas cuestiones las tocaremos de nuevo más adelante.

Si el espacio K_n es degenerado, entonces, en el caso general, la existencia de una base seudoortogonal no sirve de garantía para que coincidan los subespacios nulos derecho e izquierdo y, por tanto, no se puede esperar que el espacio se descomponga en la suma ortogonal de sus subespacios. Pero el saber la base seudoortogonal permite construir con eficacia la descomposición del espacio en la suma directa (99.1).

Supongamos que en el espacio K_n de rango r existe una base seudoortogonal $e_1, e_2, \ldots, e_n$. Convengamos en considerar que los vectores $e_1, \ldots, e_r$ son no isótropos, mientras que $e_{r+1}, \ldots, e_n$ son vectores isótropos. En la base seudoortogonal los vectores isótropos son ortogonales a la izquierda respecto a todos los vectores de la base y, consecuentemente, ortogonales a la izquierda a todos los vectores del espacio K_n. Pero esto significa que los vectores isótropos de la base seudoortogonal forman una base del subespacio nulo izquierdo ${}^{\perp}K_n$. Designemos con L la cápsula lineal de los vectores $e_1, \ldots, e_r$. De acuerdo con el corolario segundo del teorema 99.3,

$$K_n = L + {}^{\perp}L = L \dot{+} {}^{\perp}K_n,$$

con la particularidad de que para ambos subespacios L y ${}^{\perp}K_n$ las bases son conocidas. Para el subespacio L la base $e_1, \ldots, e_r$ será seudoortogonal.

Así pues, el estudio de todos los espacios degenerados con una base seudoortogonal se reduce al estudio conjunto de los subespacios regulares con una base seudoortogonal y los subespacios, donde el producto escalar es igual a cero.

Todo vector de K_n puede ser representado de modo único en forma de la suma $x = u + v$, donde $u \in L$, $v \in {}^{\perp}K_n$. Si para el vector x escribimos la descomposición (100.2), con miras a determinar los coeficientes α_j, obtendremos otra vez el sistema del tipo (100.3), pero ya no con una matriz triangular izquierda regular, sino con una matriz trapezoidal izquierda. No obstante, haciendo uso de este sistema, se pueden determinar los primeros coeficientes $\alpha_1, \ldots, \alpha_r$ y llegamos a que

$$u = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r,$$

es decir, la proyección del vector x sobre el subespacio L se halla por completo, si sólo se sabe la base seudoortogonal en L. Nuevamente $v = x - u$, y tampoco podemos hallar la descomposición del vector v según los vectores $e_{r+1}, \ldots, e_n$, recurriendo al producto escalar.

La base seudoortogonal es un tipo bastante general de base, puesto que existe casi en todos los espacios. Como ya sabemos, no existe sólo en los espacios bilineales métricos ordinarios con la forma antisimétrica (x, y). Para éstos últimos espacios el tipo más cómodo de la base es evidente y es, por supuesto, la base canónica de la matriz de Gram. Hablando en general, se puede introducir un tipo de la base que cubra todos los tipos de base considerados arriba y que exista en cualquier espacio dotado de un producto escalar. Sin embargo, esta introducción ofrece pocos hechos nuevos y por ahora no nos detendremos en este problema.

Además de una base con tales o cuales relaciones de ortogonalidad entre sus vectores, nos encontraremos a veces con los pares de bases análogas.

Una base $f_1, f_2, \ldots, f_n$ se llama *base dual izquierda* (*derecha*) para la base $e_1, e_2, \ldots, e_n$, si $(f_i, e_j) = 0$ ($(e_j, f_i) = 0$) para $i \neq j$, y, en este caso, (f_i, e_i) ((e_i, f_i)) es igual a 1 ó 0 para cualquier valor de i.

Una base $f_1, f_2, \ldots, f_n$ se llama *seudodual izquierda* (*derecha*) para la base $e_1, e_2, \ldots, e_n$, si $(f_i, e_j) = 0$ ($(e_j, f_i) = 0$) para todo

$j < i$ cuando $(f_i, e_i) = 1$ $((e_i, f_i) = 1)$ y para todo j cuando $(f_i, e_i) = 0$ $((e_i, f_i) = 0)$.

Es fácil ver que la matriz de la forma bilineal (x, y) en un par de bases duales es diagonal y en un par de bases seudoduales, trapezoidal derecha (izquierda). Las cuestiones de existencia y construcción de las bases duales y seudoduales están estrechamente relacionadas con las transformaciones equivalentes (91.4) de la matriz de la forma bilineal (x, y), como también con la descomposición de dicha matriz en factores. Recurriremos a la investigación detallada de las bases de tal tipo sólo cuando sea necesario. Aquí nos limitaremos sólo a una breve exposición de las cuestiones citadas.

TEOREMA 100.2. *En todo espacio regular cada base tiene las bases duales derecha e izquierda y éstas son únicas.*

DEMOSTRACION. Consideremos una base $e_1, e_2, \ldots, e_n$ en el espacio regular K_n y sea G_e la matriz de la forma bilineal (x, y) en dicha base. Según (91.4), el problema de búsqueda de la base dual izquierda (derecha) para $e_1, e_2, \ldots, e_n$ es equivalente a la definición de la matriz P (Q), para la cual $P'G_e$ (G_eQ) será una matriz unidad. Entonces, P (Q) será una matriz de la transformación de coordenadas al pasar de la base $e_1, e_2, \ldots, e_n$ a una base dual. Puesto que el espacio es regular, la matriz G_e también será regular y existe la única solución: $P = G_e'^{-1}$ $(Q = G_e^{-1})$.

COROLARIO. *En todo espacio regular cualquier base tiene las bases duales izquierda y derecha.*

Efectivamente, cada base dual izquierda (derecha) es a la vez una base seudodual izquierda (derecha).

Tomando en consideración la expresión para la matriz de la forma bilineal (x, y), es fácil establecer que si en un espacio regular se pasa de una base dual izquierda (derecha) a otra base que cuenta con una matriz triangular izquierda de la transformación de coordenadas cuyos elementos diagonales son unidades, entonces la base nueva será seudodual izquierda (derecha).

Ejercicios.

1. Sea simétrico un producto escalar. En el caso de las bases no ortogonales ¿se án $e_1, e_2, \ldots, e_n$ invariantes del número de vectores que tienen valores nulos, positivos y negativos de las magnitudes (e_i, e_i)?

2. ¿De qué modo se puede convertir un espacio lineal complejo o real en un espacio bilineal métrico dotado de un producto escalar simétrico en el cual se han prefijado el rango y la signatura?

3. La base ortogonal en un espacio regular no tiene vectores isótropos. ¿Podrá existir en tal espacio una base de los vectores isótropos?

4. Demuéstrese que una proyección ortogonal y una perpendicular, siendo funciones de los vectores de un espacio bilineal métrico, son operadores lineales.

5. ¿Qué forma tiene la matriz de Gram para una base seudoortogonal, si los subespacios nulos derecho e izquierdo coinciden?

6. Demuéstrese que en todo espacio bilineal métrico ordinario o hermitiano existe una base en la cual la matriz de Gram es triangular celular derecha cuyas células en la diagonal son de primero y segundo órdenes.

7. ¿De qué modo se hallan los coeficientes de la descomposición de un vector según la base, para la cual se conoce alguna base dual o seudodual?

8. Demuéstrese que en un espacio regular la matriz de la transformación de coordenadas al pasar de una base, seudodual respecto de la dada, a cualquier otra base seudodual de la misma denominación es triangular izquierda.

§ 101. Operadores y formas bilineales

Si en un espacio bilineal métrico ordinario o hermitiano actúa un operador lineal, entonces, desde luego, todos los resultados obtenidos anteriormente respecto de los operadores en un espacio real o complejo siguen siendo válidos. Por eso estudiaremos aquí sólo propiedades adicionales de los operadores relacionadas con la presencia en el espacio de un producto escalar.

Uno de los objetos más importantes es el operador conjugado. En los espacios euclídeo y unitario el operador conjugado se introducía mediante un producto escalar, mas en la investigación de sus propiedades se usaba ampliamente el hecho de existencia en el espacio de una base ortonormalizada. Ahora no podemos seguir este camino, pues en un espacio bilineal métrico general puede no haber ninguna base ortogonal. Nuestras investigaciones se realizarán en un espacio bilineal métrico hermitiano. Los cambios para el espacio bilineal métrico ordinario son muy simples.

Un operador A^* ($^*A$) que actúa en el espacio bilineal métrico hermitiano K_n se llama *conjugado derecho* (*izquierdo*) para el operador A que actúa en K_n, si para cualesquiera vectores $x, y \in K_n$ se verifica la igualdad

$$(Ax, y) = (x, A^*y) \qquad ((x, Ay) = (^*Ax, y)). \qquad (101.1)$$

Tomemos una base arbitraria $e_1, e_2, \ldots, e_n$ en K_n y sea G_e la matriz de Gram para dicha base. Designemos con A_e la matriz del operador A en la base $e_1, e_2, \ldots, e_n$ y mediante A_e^* y $^*A_e$, las matrices de los operadores A^* y $^*A$, siempre que existan.

TEOREMA 101.1. *Para todo operador lineal A que actúa en un espacio bilineal métrico hermitiano regular existen los únicos operadores conjugados A^* y $^*A$, con la particularidad de que*

$$A_e^* = \overline{G}_e^{-1}\overline{A}_e'\overline{G}_e, \qquad ^*A_e = G_e^{-1'}\overline{A}_e'G_e'. \qquad (101.2)$$

DEMOSTRACION. Si el operador A^* existe, entonces, con arreglo a la fórmula (101.1) y teniendo en cuenta las anotaciones matriciales del tipo (61.2), (91.7), resulta

$$(Ax, y) = (Ax)_e' G_e \overline{y}_e = x_e' (A_e' G_e) \overline{y}_e,$$

$$(x, A^*y) = x_e' G_e (\overline{A^*y})_e = x_e' (G_e \overline{A}_e^*) \overline{y}_e.$$

Los segundos miembros de estas correlaciones deben coincidir para todos los vectores x_e, y_e, por lo cual $A'_e G_e = G_e \overline{A}^*_e$, de donde se desprende la primera igualdad de (101.2). Por analogía,

$$(x, Ay) = x'_e G_e (\overline{Ay})_e = x'_e (G_e \overline{A}_e) \overline{y}_e,$$
$$({}^*Ax, y) = ({}^*Ax)'_e G_e \overline{y}_e = x'_e ({}^*A'_e G_e) \overline{y}_e$$

y, por eso, $G_e \overline{A}_e = {}^*A'_e G_e$ y obtenemos la segunda igualdad (101.2).

Las igualdades (101.2) significan que si los operadores conjugados existen, son únicos. Ahora convengamos en considerar estas igualdades como forma para definir los operadores conjugados derecho e izquierdo. Es fácil comprobar inmediatamente que los operadores construidos de tal modo son lineales y satisfacen las correlaciones (101.1).

COROLARIO. *Si una forma bilineal hermitiana (x, y) es simétrica o antisimétrica, los operadores conjugados derecho e izquierdo coinciden.*

En efecto, en estos casos las igualdades $G_e = \pm \overline{G}'_e$ se verifican, cualquiera que sea la base $e_1, e_2, \ldots, e_n$. En concordancia con (101.2), concluimos que $A^*_e = {}^*A_e$.

De este corolario se deduce que los operadores conjugados derecho e izquierdo coinciden en un espacio unitario. Se puede establecer también este hecho de otro modo. Si en un espacio unitario tomamos la base ortonormalizada $e_1, e_2, \ldots, e_n$, para ésta tiene lugar la igualdad $G_e = E$ y obtenemos las igualdades bien conocidas $A^*_e = {}^*A_e = \overline{A}'_e$.

Los operadores conjugados están ligados con el operador A por unas correlaciones determinadas. Demos a conocer algunas de ellas, por ejemplo, para el operador conjugado derecho:

$$(A+B)^* = A^* + B^*,$$
$$(\alpha A)^* = \overline{\alpha} A^*, \qquad (AB^*) = B^* A^*, \qquad (A^*)^{-1} = (A^{-1})^*. \quad (101.3)$$

Para el operador conjugado izquierdo las correlaciones son análogas. Todas las correlaciones se demuestran siguiendo el mismo esquema, empleándose las representaciones (101.2) para las matrices de los operadores conjugados. Por esto nuestro intento es sólo demostrar la validez de la última propiedad. Se tiene

$$(A^*_e)^{-1} = \overline{G}_e^{-1} (\overline{A}'_e)^{-1} (\overline{G}_e^{-1})^{-1} = \overline{G}_e^{-1} (\overline{A_e^{-1}})' \overline{G}_e = (A_e^{-1})^*.$$

Comparando las fórmulas (75.4), (101.3) se puede advertir la ausencia en (101.3) del análogo de la primera correspondencia (75.4). Ahora dicha correspondencia tiene la forma:

$$({}^*A)^* = {}^*(A^*) = A. \quad (101.4)$$

Con el objeto de demostrar que es verídico, recurrimos otra vez a las representaciones (101.2) y obtenemos

$$({}^{*}A_e)^{*} = \overline{G}_e^{-1}\,(\overline{{}^{*}A_e})'\,\overline{G}_e = \overline{G}_e^{-1}\,(\overline{G_e^{-1}\overline{A}_e'G_e'})'\,\overline{G}_e = \overline{G}_e^{-1}\overline{G}_e A_e \overline{G}_e^{-1}\overline{G}_e = A_e,$$

$${}^{*}(A_e^{*}) = G_e^{-1'}\,(\overline{A_e^{*}})'\,G_e' = G_e^{-1'}\,(\overline{\overline{G}_e^{-1}\overline{A}_e'\overline{G}_e})'\,G_e' = G_e^{-1'}G_e'A_eG_e^{-1'}G_e' = A_e,$$

es decir, las correlaciones (101.4) son realmente justas.

TEOREMA 101.2. *Si en un espacio bilineal métrico hermitiano regular el operador A tiene en cierta base la matriz J, entonces en la base dual derecha (izquierda) el operador A^{*} (${}^{*}A$) cuenta con la matriz J^{*}.*

DEMOSTRACION. Supongamos que el operador A tiene una matriz J en la base $e_1, e_2, \ldots e_n$. Examinemos la base dual derecha $f_1, f_2, \ldots, f_n$. Designemos mediante G_e, G_f y $G_{ef} = E$ las matrices de la forma bilineal (x, y) en las bases correspondientes. Si P es la matriz de la transformación de coordenadas, al pasar de la primera base a la segunda, resulta

$$G_e = G_{ef}\overline{P}^{-1} = \overline{P}^{-1}, \qquad G_f = P'G_{ef} = P',$$

y luego, al tomar en consideración (63.7), (101.2), obtenemos

$$A_f^{*} = \overline{G}_f^{-1}\overline{A}_f'\overline{G}_f = \overline{G}_f^{-1}\,(\overline{P^{-1}JP})'\,\overline{G}_f = \overline{G}_f^{-1}\overline{G}_f\overline{J}'\overline{G}_f^{-1}\overline{G}_f = J^{*}.$$

En cambio, si el operador A tiene la matriz J en la base $f_1, f_2, \ldots, f_n$, entonces, para dicha base, la base $e_1, e_2, \ldots, e_n$ será dual izquierda y ahora tenemos

$${}^{*}A_e = G_e^{-1'}\,\overline{A}_e'G_e' = G_e^{-1'}\,\overline{(PA_fP^{-1})}'\,G_e' = G_e^{-1'}G_e'\overline{J}'G_e^{-1'}G_e' = J^{*}.$$

El teorema demostrado es de la misma significación al investigar los operadores conjugados en espacios bilineales métricos hermitianos que el teorema 75.2 en los espacios unitarios. En particular, de este teorema se infiere que los operadores conjugados derecho e izquierdo A^{*} y ${}^{*}A$ poseen los mismos valores propios complejos conjugados respecto de los valores propios del operador A y que los operadores conjugados derecho e izquierdo A^{*} y ${}^{*}A$ son de estructura simple, siempre que tenga la misma estructura el operador A, etc.

Además del producto escalar (x, y), en un espacio bilineal métrico hermitiano pueden definirse también otras formas bilineales hermitianas. Consideraremos, por ejemplo, las funciones del tipo (Ax, y) y (x, Ay), donde A es un operador lineal arbitrario. No es difícil convencerse de que estas funciones son unas formas bilineales hermitianas. En cualquier espacio regular K_n los diferentes operadores definen formas distintas. En efecto, si A, B son operadores diferentes, por lo menos para un solo vector x se verifica la desigualdad $Ax \neq Bx$. Supongamos que para todo $y \in K_n$ se verifica la igualdad $(Ax, y) = (Bx, y)$. De aquí se desprende que $((A - B)\,x, y) = 0$ para todo $y \in K_n$, es decir, $(A - B)\,x \in {}^{\perp}K_n$. Pero, en un espacio regular el subespacio ${}^{\perp}K_n$ sólo se compone del vector nulo y, por eso, $Ax = Bx$.

TEOREMA 101.3. *En un espacio bilineal métrico hermitiano regular K_n cualquier forma bilineal hermitiana $\varphi(x, y)$ puede ser representada de un modo único como la expresión*

$$\varphi(x, y) = (Ax, y) = (x, By),$$

donde A, B son ciertos operadores lineales que actúan en K_n.

DEMOSTRACIÓN. Elijamos en el espacio K_n una base $e_1, e_2, \ldots, e_n$ y sea G_e la matriz de Gram en dicha base y Φ_e, la matriz de la forma $\varphi(x, y)$. Tenemos

$$\varphi(x, y) = x'_e\Phi_e\overline{y}_e = x'_e\Phi_e G_e^{-1} G_e\overline{y}_e =$$
$$= (G_e^{-1'}\Phi'_e x_e)' G_e\overline{y}_e = x'_e G_e G_e^{-1}\Phi_e\overline{y}_e = x'_e G_e (\overline{\overline{G}_e^{-1}\overline{\Phi}_e y_e}).$$

Ahora las matrices A_e, B_e de los operadores buscados se determinan mediante las igualdades

$$A_e = G_e^{-1'}\Phi'_e, \qquad B_e = \overline{G}_e^{-1}\overline{\Phi}_e. \tag{101.5}$$

La unicidad de los operadores A, B se ha demostrado antes.

Un operador conjugado se define en términos del producto escalar. Por esto, si en un espacio lineal se introducen productos escalares diferentes, entonces un mismo operador lineal tendrá diferentes operadores conjugados. Supongamos que en un espacio lineal, junto con el producto escalar prefijado por la forma bilineal (x, y), se introducen, además, unos productos escalares prefijados mediante las formas (Mx, y), (x, My). Indiquemos con M, al poner este índice abajo a la izquierda (a la derecha), los operadores conjugados referentes al producto escalar (Mx, y) $((x, My))$.

TEOREMA 101.4. *Para cualquier operador A y un operador regular M tienen lugar las correlaciones*

$$\begin{aligned} {}_M A^* &= (MAM^{-1})^*, & {}^*_M A &= M^{-1}({}^*A)M, \\ A^*_M &= M^{-1}A^*M, & {}^*A_M &= {}^*_{\cdot}(MAM^{-1}). \end{aligned} \tag{101.6}$$

DEMOSTRACIÓN. Elijamos una base cualquiera $e_1, e_2, \ldots, e_n$ y sean G_e y M_e las matrices de la forma bilineal (x, y) y del operador M en esta base, respectivamente. De acuerdo con (101.5), la matriz de la forma bilineal (Mx, y) es igual a $M'G_e$. Ahora, de conformidad con (101.2) encontramos

$${}_M A^*_e = \overline{(M'_eG_e)^{-1}}\,\overline{A}'_e(\overline{M'_eG_e}) = \overline{G}_e^{-1}(\overline{M}_e^{-1'}\overline{A}'_e\overline{M}'_e)\overline{G}_e =$$
$$= \overline{G}_e^{-1}(\overline{M_eA_eM_e^{-1}})'\,\overline{G}_e = (MAM^{-1})^*_e.$$

$${}^*_M A_e = (M'_eG_e)^{-1'}\,\overline{A}'_e(M'G_e)' = M_e^{-1'}G_e^{-1'}\overline{A}'_eG'_eM =$$
$$= M_e^{-1}(G_e^{-1'}\overline{A}'_eG'_e)M_e = M_e^{-1}({}^*A_e)M_e.$$

Las igualdades matriciales obtenidas demuestran la validez del primer grupo de igualdades operacionales (101.6). El segundo grupo se desprende evidentemente del primero, si se toman en consideración la igualdad $(x, My) = (^*Mx, y)$ y las correlaciones (101.2).

En un espacio bilineal métrico hermitiano se consideran diferentes tipos de los operadores. El operador A se llama *hermitiano* o *autoconjugado*, si para cualesquiera $x, y \in K_n$

$$(Ax, y) = (x, Ay),$$

y *antihermitiano* o *anticonjugado*, si

$$(Ax, y) = -(x, Ay).$$

De aquí provienen las igualdades respectivas

$$A = A^* = {}^*A, \qquad A = -A^* = -{}^*A.$$

El operador A se denomina *isométrico*, si para cualesquiera $x, y \in K_n$ se tiene

$$(Ax, Ay) = (x, y).$$

Esto nos conduce a las igualdades

$${}^*AA = A^*A.$$

En un espacio bilineal métrico ordinario los análogos de los operadores hermitiano y antihermitiano se llaman *simétrico* y *antisimétrico*, respectivamente. En lo que sigue nos encontraremos más de una vez con los operadores que se definen mediante la igualdad

$$A^* = \alpha E + \beta A \tag{101.7}$$

para ciertos números α, β.

No todas las propiedades de los operadores del tipo especial, ni mucho menos, se pueden transferir de un espacio unitario a un espacio bilineal métrico hermitiano, aunque tienen algo en común. Aquí pasamos por alto las investigaciones de todas estas cuestiones.

Ejercicios.

1. ¿De qué modo están ligados entre sí los polinomios característicos de los operadores A, A^*, ${}^*A$?

2. Supongamos que el subespacio L es invariante respecto del operador A. Demuéstrese que el subespacio $L\perp$ $(\perp L)$ es invariante respecto del operador A^* $({}^*A)$.

3. Demuéstrese que cualquier vector propio del operador A, correspondiente al valor propio λ, es ortogonal a la izquierda (a la derecha) a todo vector propio del operador A^* $({}^*A)$ que corresponde al valor propio $\mu \neq \bar{\lambda}$.

4. Demuéstrese que cualquier vector rádical del operador A, correspondiente al valor propio λ, es ortogonal a la izquierda (a la derecha) a todo vector radical del operador A^* $({}^*A)$ que corresponde al valor propio $\mu \neq \bar{\lambda}$.

5. Demuéstrese que los valores propios de un operador hermitiano (antihermitiano), correspondientes a los vectores propios no isótropos, son reales (imaginarios puros).

6. Demuéstrese que los módulos de los valores propios de un operador isométrico, correspondientes a los vectores propios no isótropos, son iguales a la unidad.

7. Supongamos que en un espacio regular el producto escalar es simétrico según Hermite. Demuéstrese que si el operador A es hermitiano (antihermitiano), entonces la forma bilineal (Ax, y) es simétrica (antisimétrica) según Hermite.

8. Supongamos que en un espacio regular el producto escalar es simétrico según Hermite. Demuéstrese que si la forma bilineal (Ax, y) es simétrica (antisimétrica) según Hermite, el operador A es hermitiano (antihermitiano).

9. ¿De qué modo cambian las afirmaciones de los ejercicios 7, 8, si el producto escalar es antisimétrico según Hermite?

10. Demuéstrese que si el operador A, que satisface la condición (101.7), tiene por lo menos dos valores propios distintos, entonces $|\beta| = 1$.

§ 102. Isomorfismo bilineal métrico

Al investigar los espacios euclídeos y unitarios hemos demostrado que existe, salvo un isomorfismo, un solo espacio de cada dimensión n. Para los espacios bilineales métricos las cosas resultan ser más complejas.

Introduzcamos el concepto de isomorfismo. Diremos que los espacios ordinarios o bilineales métricos hermitianos sobre un mismo campo numérico son *isomorfos*, si son isomorfos como espacios lineales, con la particularidad de que los productos escalares de los pares de vectores correspondientes son iguales entre sí.

De esta definición se deduce que en los espacios isomorfos las matrices de Gram de los sistemas de vectores correspondientes coinciden. La afirmación recíproca es también justa. Si en los espacios bilineales métricos sobre un campo numérico común existen bases con matrices de Gram coincidentes, estos espacios son isomorfos. Efectivamente, al establecer la correspondencia entre las bases con matrices iguales de Gram, aseguramos la coincidencia de los productos escalares para cualesquiera pares de vectores de las bases y, consecuentemente, para cualesquiera pares de vectores.

TEOREMA 102.1. *Los espacios bilineales métricos ordinarios (hermitianosos) sobre un mismo campo numérico son isomorfos, si y sólo si, las matrices de Gram de las bases arbitrarias de estos espacios son congruentes (congruentes según Hermite).*

DEMOSTRACIÓN. NECESIDAD. Las matrices de Gram de todas las bases de un mismo espacio son congruentes, y coinciden en las bases correspondientes de unos espacios diferentes. Por ser transitiva la relacion de congruencia, las matrices de Gram para las bases arbitrarias de los espacios isomorfos serán congruentes.

SUFICIENCIA. Si las matrices de Gram para las bases arbitrarias de los espacios bilineales métricos son congruentes, existen en diferentes espacios unas bases, donde las matrices de Gram coinciden. Mas, en este caso los espacios son isomorfos.

El teorema demostrado dice que el problema de clasificación de los espacios bilineales métricos es equivalente al problema de clasificación de las formas bilineales, salvo la congruencia. Examinaremos algunas clases de los espacios bilineales métricos.

Un espacio bilineal métrico real K_n se llama *seudoeuclidiano*, si el producto escalar viene dado por una forma bilineal simétrica regular.

Para una base arbitraria de un espacio seudoeuclidiano la matriz de Gram es real simétrica y, como ya sabemos, congruente de la matriz diagonal con los elementos ± 1. Esto significa que en todo espacio seudoeuclidiano existe una base en la que el producto escalar (x, y) de los vectores x, y, cuyas coordenadas son $\xi_1, \ldots, \xi_n$ y $\eta_1, \ldots, \eta_n$, se da mediante la fórmula

$$(x, y) = \xi_1\eta_1 + \ldots + \xi_s\eta_s - \xi_{s+1}\eta_{s+1} - \ldots - \xi_n\eta_n.$$

Los espacios seudoeuclidianos se determinan, salvo un isomorfismo, por sus dos características: la dimensión y la signatura, los índices positivo y negativo, etc. Entre los espacios seudoeuclidianos es de mayor interés para la física un espacio cuadridimensional de índice positivo igual a uno. Este es el así llamado espacio *de Minkowski*. Representa en sí un espacio de sucesos de la teoría especial de la relatividad.

Un espacio bilineal métrico real K_n se denomina *simplicial*, si el producto escalar viene dado por una forma bilineal antisimétrica regular.

La matriz de Gram para cualquier espacio simplicial es antisimétrica y, debido a esta circunstancia, es congruente de la matriz celular diagonal con células del tipo $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Por esta razón la dimensión de un espacio simplicial es siempre par y existe, salvo un isomorfismo, un solo espacio simplicial de dimensión par prefijada. En tal espacio existe una base en la que el producto escalar de los vectores x, y con las coordenadas $\xi_1, \ldots, \xi_n$ y $\eta_1, \ldots, \eta_n$ es de la orma

$$(x, y) = \xi_1\eta_2 - \xi_2\eta_1 + \ldots + \xi_{n-1}\eta_n - \xi_n\eta_{n-1}.$$

Un espacio bilineal métrico complejo K_n se llama *euclidiano complejo*, si el producto escalar viene dado por una forma bilineal simétrica regular.

Para cualquier base la matriz de Gram es simétrica compleja y congruente de la matriz unidad. Existe, salvo un isomorfismo, un solo espacio euclidiano complejo de cada dimensión. En todo espacio euclidiano complejo existe una base en la que el producto escalar de los vectores x, y es de la forma

$$(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n.$$

Un espacio bilineal métrico hermitiano complejo se denomina *seudounitario*, si el producto escalar viene dado por una forma bilineal simétrica hermitiana regular.

La matriz de Gram para todo espacio seudounitario es hermitiana. Es congruente según Hermite de la matriz diagonal real con los elementos ± 1. Por esta razón, existe siempre una base en la que el producto escalar de los vectores x, y tiene por expresión

$$(x, y) = \xi_1 \overline{\eta}_1 + \ldots + \xi_s \overline{\eta}_s - \xi_{s+1} \overline{\eta}_{s+1} - \ldots - \xi_n \overline{\eta}_n,$$

donde $\xi_1, \ldots, \xi_n$ y $\eta_1, \ldots, \eta_n$ son las coordenadas de los vectores x, y. Esta vez también, un espacio seudounitario se determina unívocamente, salvo un isomorfismo, mediante dos características suyas: la dimensión y la signatura, los índices positivo y negativo, etc.

Ejercicios

1. Demuéstrese que en los espacios isomorfos, a las bases ortogonales (seudoortogonales, duales, seudoduales) les corresponden unas bases ortogonales (seudoortogonales, duales, seudoduales).

2. Demuéstrese que en los espacios isomorfos, a los subespacios regulares les corresponden unos subespacios regulares.

3. Demuéstrese que en los espacios isomorfos una perpendicular y una proyección pasan a ser una perpendicular y una proyección, respectivamente.

4. Demuéstrese que en los espacios isomorfos los determinantes de Gram de los sistemas correspondientes de vectores son iguales.

CAPÍTULO 13 FORMAS BILINEALES EN LOS PROCESOS DE CÁLCULO

§ 103. Procesos de ortogonalización

Uno de los conceptos más importantes relacionados con cualquier espacio bilineal métrico es el de ortogonalización. Ya nos convencimos más de una vez cuán importante es el papel que desempeñan los sistemas ortogonales de vectores y sobre todo las bases ortogonales en el estudio de los espacios euclidianos y unitarios. No es menor el papel que las bases con vectores ortogonales desempeñan en otros espacios. No obstante, la mayor parte de nuestros razonamientos ha sido asociada hasta ahora con la demostración de la existencia de tales sistemas y no con los procesos de su construcción. Cierta excepción representa solamente el método general de la transformación de matrices de las formas bilineales a la forma canónica y la construcción de las bases canónicas vinculada con la transformación citada. En vista de que los sistemas ortogonales, seudoortogonales y otros análogos son muy esenciales en la construcción de los más diversos algoritmos de cálculo, examinemos ahora un proceso general que tiene por objeto la construcción de semejantes sistemas en un espacio bilineal métrico.

Supongamos que en un espacio lineal complejo K_n viene dado, con ayuda de una forma bilineal hermitiana regular, el producto escalar (x, y). Consideraremos una base $e_1, e_2, \ldots, e_n$ y trataremos de construir otra base $f_1, f_2, \ldots, f_n$ que posea las siguientes propiedades:

1) las cápsulas lineales L_k de los vectores $e_1, e_2, \ldots, e_k$ y $f_1, f_2, \ldots, f_k$ coinciden para todo $k \geqslant 1$,

2) la base $f_1, \ldots, f_n$ es seudoortogonal. Supongamos que $(e_1, e_1) \neq 0$ y hagamos $f_1 = e_1$. Sea ya construido el sistema de los vectores seudoortogonales $f_1, \ldots, f_k$, con la particularidad de que las cápsulas lineales de estos vectores y de los vectores $e_1, \ldots, e_k$ coinciden y $(f_i, f_i) \neq 0$ para $1 \leqslant i \leqslant k$. Buscaremos el vector f_{k+1} en la forma

$$f_{k+1} = e_{k+1} + \sum_{i=1}^{k} \alpha_{i,\, k+1} f_i. \qquad (103.1)$$

donde $\alpha_{1,k+1}, \ldots, \alpha_{k,k+1}$ son unos coeficientes desconocidos. Las condiciones de ortogonalidad del vector f_{k+1} a la izquierda respecto a los vectores $f_1, \ldots, f_k$ ofrecen, para determinar $\alpha_{1,k+1}, \ldots, \alpha_{k,k+1}$, un sistema de ecuaciones algebraicas lineales

$$\begin{aligned}
&\alpha_{1,k+1}(f_1, f_1) && = -(e_{k+1}, f_1),\\
&\alpha_{1,k+1}(f_1, f_2) + \alpha_{2,k+1}(f_2, f_2) && = -(e_{k+1}, f_2),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots && \\
&\alpha_{1,k+1}(f_1, f_k) + \alpha_{2,k+1}(f_2, f_k) + \ldots + \alpha_{k,k+1}(f_k, f_k) && = -(e_{k+1}, f_k).
\end{aligned} \tag{103.2}$$

La matriz de este sistema es triangular izquierda. Por hipótesis, sus elementos diagonales se difieren de cero, debido a lo que el sistema (103.2) tiene la única solución. Está claro que el vector f_{k+1}, construido de tal modo, junto con los vectores $f_1, \ldots, f_k$ forman un sistema seudoortogonal y su cápsula lineal coincide con la de los vectores $e_1, \ldots, e_{k+1}$. El sistema de vectores $f_1, \ldots, f_{k+1}$ es linealmente independiente, puesto que linealmente independiente es el sistema $f_1, \ldots, f_k, e_{k+1}$.

Seguimos adelante el proceso. Si resulta que las magnitudes (f_i, f_i) son distintas de cero para i cualquiera, entonces el sistema obtenido de vectores $f_1, \ldots, f_n$ será precisamente la base seudoortogonal buscada. Desde luego, podemos ahora normalizar los vectores $f_1, \ldots, f_n$ y obtener una base ortonormalizada.

De la fórmula (103.1) se desprende un corolario de importancia. Escribamos la igualdad (103.1) bajo la forma

$$e_{k+1} = \left(-\sum_{i=1}^{k} \alpha_{i,k+1} f_i\right) + f_{k+1}.$$

El vector entre paréntesis pertenece a L_k, el vector f_{k+1} pertenece a ${}^{\perp}L_k$ por construcción, por lo cual la solución de los sistemas (103.2) nos da, en realidad, la descomposición de cada vector e_{k+1} en una proyección y la perpendicular izquierda respecto del subespacio L_k.

El proceso descrito se simplifica de modo considerable, si el producto escalar viene dado por una forma bilineal simétrica hermitiana. En tal caso las condiciones $(f_i, f_j) = 0$ para $j < i$ llevan consigo el cumplimiento de las condiciones $(f_i, f_j) = 0$ para $j \neq i$. Por ello, el sistema (103.2) se convierte en un sistema con una matriz diagonal y se tiene

$$\alpha_{i,k+1} = -\frac{(e_{k+1}, f_i)}{(f_i, f_i)}$$

para todo i. La base construida $f_1, f_2, \ldots, f_n$ será no sólo seudoortogonal, sino también ortogonal.

La única causa por la cual puede obstacularizarse la construcción de la base seudoortogonal $f_1, \ldots, f_n$, a partir de la base $e_1, \ldots, e_n$, consiste en la anulación de uno de los productos escalares (f_i, f_i), $i < n$. Tal situación se llamará degenerada. La situación degenerada no viene a ciencia cierta, si la forma cuadrática (x, x) no tiene vectores isótropos, por ejemplo, si esta última es estrictamente de signo constante. Efectivamente, en este caso la igualdad $(f_i, f_i) = 0$ es posible sólo cuando $f_i = 0$. Mas, $f_i \neq 0$ para i cualquiera, puesto que los vectores $f_1, \ldots, f_i$ son linealmente independientes. Por consiguiente, ahora el proceso es realizable, cualquiera que sea la elección de la base $e_1, \ldots, e_n$.

Existen muchos problemas en los cuales no hace falta conservar los lazos de la nueva base $f_1, \ldots, f_n$ con la base inicial $e_1, \ldots, e_n$, puesto que se necesita sólo construir en el espacio una base seudoortogonal. En este caso, cada vez que aparezca la igualdad $(f_i, f_i) = 0$, se debe sustituir el vector e_i por otro y calcular de nuevo el vector f_i, repitiendo este procedimiento hasta que se cumpla la condición $(f_i, f_i) \neq 0$. Los vectores $f_1, \ldots, f_{i-1}$ quedan invariables.

Un vector necesario para sustituir e_i siempre existe. Supongamos que la igualdad $(f_i, f_i) = 0$ se verifica para cualquier vector e_i. Puesto que el vector f_i es ortogonal a la izquierda respecto a los vectores $e_1, \ldots, e_{i-1}$, esto significa que el subespacio ${}^{\perp}L_{i-1}$ está compuesto sólo por los vectores isótropos y el vector nulo. Mas, el subespacio L_{i-1} es regular, por lo cual ${}^{\perp}L_{i-1} = {}^{\perp}K_n$. La última igualdad no puede tener lugar para $i - 1 < n$, puesto que, por ser K_n regular, el subespacio ${}^{\perp}K_n$ sólo consta del vector nulo.

Al proceder del mismo modo, podemos construir también una base que sea seudodual para la dada. Supongamos otra vez que $e_1, e_2, \ldots, e_n$ es la base dada y es necesario construir una base que sea seudodual para la base dada, por ejemplo, la izquierda. Tomemos una base más: $q_1, q_2, \ldots, q_n$. Sea $(q_1, e_1) \neq 0$ y pongamos $t_1 = q_1$. Admitamos que ya se ha construido un sistema de vectores $t_1, \ldots, t_k$ tales que su cápsula lineal coincide con la cápsula lineal de los vectores $q_1, \ldots, q_k$ y se cumplen las condiciones $(t_i, e_i) \neq 0$ para $1 \leqslant i \leqslant k$ y $(t_i, e_j) = 0$ para $j < i$. Buscaremos el vector t_{k+1} en forma

$$t_{k+1} = q_{k+1} + \sum_{i=1}^{k} \beta_{i,\,k+1} t_i, \tag{103.3}$$

donde $\beta_{1,k+1}, \ldots, \beta_{k,\,k+1}$ son unos coeficientes desconocidos. La condición de ortogonalidad del vector t_{k+1} a la izquierda respecto de los vectores $e_1, \ldots, e_k$ nos ofrece de nuevo, para determinar $\beta_{1,\,k+1}, \ldots, \beta_{k,\,k+1}$, un sistema de ecuaciones algebraicas lineales provisto de una matriz triangular izquierda:

$$\begin{aligned}
&\beta_{1,\,k+1}(t_1,\, e_1) &&= -(q_{k+1},\, e_1)\\
&\beta_{1,\,k+1}(t_1,\, e_2)+\beta_{2,\,k+1}(t_2,\, e_2) &&= -(q_{k+1},\, e_2)\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&\beta_{1,\,k+1}(t_1,\, e_k)+\beta_{2,\,k+1}(t_2,\, e_k)+\ldots+\beta_{k,\,k+1}(t_k,\, e_k) &&= -(q_{k+1},\, e_k).
\end{aligned} \tag{103.4}$$

De acuerdo con la suposición respecto a los elementos diagonales, el sistema tiene la solución única. Si, al continuar el proceso, resulta que las magnitudes $(t_i,\, e_i)$ son distintas de cero para i cualquiera, entonces el sistema obtenido de vectores será, siendo normalizado adecuadamente, la base seudodual para $e_1,\, e_2,\, \ldots,\, e_n$. Cabe notar que esta vez el proceso no se simplifica, si el producto escalar viene dado mediante una forma bilineal simétrica. El empleo de la base auxiliar $q_1,\, q_2,\, \ldots,\, q_n$ permite evitar la degeneración del proceso, sustituyendo en el momento apropiado uno de los vectores q_i y repitiendo los cálculos del vector t_i. En el caso dado tampoco cambian los vectores $t_1,\, \ldots,\, t_{i-1}$.

Todos los procesos descritos y los procesos análogos se llamarán, con mayor frecuencia independientemente de su contenido concreto, *procesos de ortogonalización*. No obstante, a veces nos veremos obligados a construir, en un mismo espacio bilineal métrico K_n, unas sucesiones de vectores, ortogonales o seudoortogonales con relación a las diferentes formas bilineales. Se considerarán sólo las formas del tipo $(Rx,\, y)$, donde R es un operador lineal en K_n. Para poder distinguir diferentes sucesiones una de la otra, diremos que se trata de la R-ortogonalización, la R-seudoortogonalización, etc.

Muchas propiedades y peculiaridades de los procesos de ortogonalización pueden establecerse considerando sus notaciones matriciales. Supongamos que un producto escalar en K_n viene dado por la forma bilineal hermitiana $(x,\, y)$. La seudoortogonalidad de la base $f_1,\, f_2,\, \ldots,\, f_n$ significa que $(f_i,\, f_j) = 0$ para $j < i$, es decir, la matriz de Gram G_f de la forma bilineal $(x,\, y)$ en la base $f_1,\, f_2,\, \ldots,\, f_n$ será triangular derecha. Conforme al proceso de construcción de la nueva base, las cápsulas lineales de los vectores $f_1,\, \ldots,\, f_k$ y $e_1,\, \ldots,\, e_k$ coinciden. Por consiguiente, al tomar en consideración (103.1), concluimos que

$$\begin{aligned}
e_1 &= f_1,\\
e_2 &= -\alpha_{1,\,2} f_1 + f_2\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
e_n &= -\alpha_{1,\,n} f_1 - \alpha_{2,\,n} f_2 - \ldots - \alpha_{n-1,\,n} f_{n-1} + f_n,
\end{aligned} \tag{103.5}$$

donde α_{ij} son precisamente aquellos coeficientes que se calculan partiendo de los sistemas (103.2). Por ello, la matriz A de la transformación de coordenadas al pasar de la nueve base $f_1, f_2, \ldots, f_n$ a $e_1, e_2, \ldots, e_n$ es triangular derecha cuyos elementos diagonales son iguales a la unidad. Dado que la matriz de la transformación de coordenadas al pasar de la base antigua a la nueva coincide con A^{-1}, se tiene

$$G_f = A^{-1'} G_e \bar{A}^{-1}.$$

De aquí proviene la igualdad

$$G_e = A' G_f \bar{A}. \tag{103.6}$$

Es fácil comprobar que la matriz $G_f\bar{A}$ es triangular derecha y sus elementos diagonales coinciden con los elementos diagonales de la matriz G_f.

Designemos con E_q (F_q) una matriz cuyas columnas están representadas por las coordenadas de los vectores $e_1, \ldots, e_n$ ($f_1, \ldots, f_n$) en la base $q_1, \ldots, q_n$. Las correlaciones (103.5) muestran que

$$E_q = F_q A, \tag{103.7}$$

y, por supuesto, además,

$$G_f = F'_q G_q \bar{F}_q. \tag{103.8}$$

Así pues, el proceso considerado de construcción de una base seudoortogonal resulta estrechamente relacionado con la descomposición de la matriz de Gram en factores triangulares y con la descomposición (103.7) en factores de la matriz de las coordenadas.

TEOREMA 103.1. *Para que el proceso* (103.1), (103.2) *de construcción de la base seudoortogonal* $f_1, f_2, \ldots, f_n$ *partiendo de la base* $e_1, e_2, \ldots, e_n$, *sea realizable en un espacio bilineal métrico regular* K_n, *es necesario y suficiente que la matriz de Gram del sistema* $e_1, e_2, \ldots, e_n$ *tenga menores principales no nulos.*

DEMOSTRACIÓN. NECESIDAD. Supongamos que el proceso es realizable, es decir, tiene lugar la correlación (103.6). La matriz G_f es regular, puesto que representa la matriz de Gram de la forma bilineal regular (x, y) para la base. Por esta razón todos sus elementos diagonales son distintos de cero. Aplicando la fórmula de Binet-Cauchy, obtenemos que para todo r

$$G_e\begin{pmatrix}1 & 2 & \ldots & r\\ 1 & 2 & \ldots & r\end{pmatrix} = A'G_f\bar{A}\begin{pmatrix}1 & 2 & \ldots & r\\ 1 & 2 & \ldots & r\end{pmatrix} = G_f\begin{pmatrix}1 & 2 & \ldots & r\\ 1 & 2 & \ldots & r\end{pmatrix} \neq 0.$$

SUFICIENCIA. Supongamos que los menores principales de la matriz de Gram G_e son distintos de cero. Por lo tanto, en concordancia con (93.1), existe una descomposición $G_e = L_e D_e U_e$, donde L_e es una matriz triangular izquierda con elementos diagonales unidad, D_e es una matriz diagonal con elementos no nulos, U_e es una matriz triangular derecha con elementos diagonales unidad. Es fácil ver que la matriz

$$G_f = \overline{U}_e^{-1'} G_e U_e^{-1} = \overline{U}_e^{-1'} L_e D_e$$

es triangular izquierda cuyos elementos diagonales coinciden con los elementos diagonales de la matriz D_e. Ahora, si a título de la matriz A tomamos la matriz triangular derecha $\overline{U}_e$, con elementos diagonales unidad, entonces para la base $f_1, f_2, \ldots, f_n$ se cumplirán las correlaciones (103.5). Precisamente esta base será construida conforme al proceso (103.1), (103.2), de lo que es fácil convencerse por comprobación inmediata.

Si el producto escalar está dado por una forma bilineal hermitiana simétrica, la matriz G_e será hermitiana, como lo será también la matriz G_f. Pero de aquí se deduce que la matriz G_f será diagonal. Este hecho ya se ha notado. Comparando (93.5), (103.6), concluimos que el proceso de ortogonalización en el caso dado coincide completamente con el de la obtención de la descomposición (93.5).

Si el espacio K_n es unitario, el proceso de ortogonalización determina no sólo la descomposición de la matriz de Gram en factores triangulares, sino también la descomposición de la matriz de las coordenadas en un producto de los factores unitario y triangular derecho. En efecto, elijamos una base ortonormalizada $q_1, q_2, \ldots, q_n$ e indiquemos con D_q una matriz diagonal formada de las longitudes de las columnas de la matriz F_q de (103.7). Ahora tenemos $E_q = (F_q D_q^{-1}) (D_q A)$. La matriz $D_q A$ es triangular derecha. Pero las matrices G_q y $G_f (D_q^{-1})^2$ son matrices unidad. Según (103.8), en el caso dado $(F_q D_q^{-1})' \overline{(F_q D_q^{-1})} = E$, es decir, la matriz $F_q D_q^{-1}$ es unitaria.

El hecho de que la base $t_1, t_2, \ldots, t_n$ es seudodual izquierda para la base $e_1, e_2, \ldots, e_n$ testimonia que para la forma bilineal (x, y), que determina el producto escalar en K_n, se cumplen las condiciones $(t_i, e_j) = 0$ para $j < i$ y $(t_i, e_j) = 1$ para cualquier i. En otras palabras, esto significa que para el par de bases $e_1, e_2, \ldots, e_n$ y $t_1, t_2, \ldots, t_n$ la matriz G_{te} de la forma bilineal (x, y) es triangular derecha con elementos diagonales unidad. De aquí concluimos que la matriz Q^{-1} de la transformación de coordenadas, al pasar de la base inicial $q_1, q_2, \ldots, q_n$ a la base $t_1, t_2, \ldots, t_n$, es triangular derecha. Sin embargo, en este caso los elementos diagonales no serán iguales a unidades, puesto que los vectores $t_1, t_2, \ldots, t_n$ se han sometido a la normalización. Se tiene

$$G_{te} = Q^{-1'} G_{qe},$$

y luego

$$G_{qe} = Q' G_{te}.$$

El proceso de construcción de la base, seudodual respecto a la dada, resulta también estrechamente ligado con la descomposición (93.1) de una matriz en factores triangulares.

TEOREMA 103.2. *Para que sea realizable el proceso de construcción* (103.3), (103.4) *de la base* $t_1, t_2, \ldots, t_n$, *izquierda seudodual para* $e_1, e_2, \ldots, e_n$, *a partir de la base* $q_1, q_2, \ldots, q_n$, *es necesario y suficiente que la matriz* G_{qe} *de la forma bilineal* (x, y) *tenga en las bases* $q_1, q_2, \ldots, q_n$ *y* $e_1, e_2, \ldots, e_n$ *los menores principales no nulos.*

La demostración de este teorema se omite aquí, puesto que es casi la repetición textual de la demostración del teorema anterior.

Subrayemos, como conclusión, que los procesos considerados de ortogonalización se extienden completamente al caso de los espacios bilineales métricos ordinarios. Cambian sólo algunos detalles relacionados con la conjugación compleja. Además, en el caso dado resulta más difícil eliminar situaciones degeneradas.

Ejercicios.

1. ¿Cuál es la interpretación geométrica del proceso de ortogonalización?
2. Demuéstrese que si el proceso de ortogonalización se aplica a un sistema linealmente dependiente $e_1, e_2, \ldots, e_n$, entonces $f_k = 0$ para cierto $k \leqslant n$.
3. Supongamos que la forma cuadrática (x, x) está privada de vectores isótropos. ¿Cómo se determina la base del sistema dado de vectores con ayuda del proceso de ortogonalización?
4. Demuéstrese que si un proceso de ortogonalización se realiza en un espacio euclídeo o en un espacio unitario, la desigualdad $|f_k| \leqslant |e_k|$ se cumple para k cualquiera, con la particularidad de que la igualdad se logra cuando, y sólo cuando, el vector e_k sea ortogonal a los vectores $e_1, \ldots, e_{k-1}$.
5. Supongamos que las coordenadas de los vectores $e_1, \ldots, e_n$ en cierta base ortonormalizada de un espacio euclídeo o unitario forman una matriz triangular. ¿Cómo varía la matriz de las coordenadas después de realizarse el proceso de ortogonalización?
6. ¿Se podrá construir una base dual con ayuda de un proceso de ortogonalización?
7. ¿Cómo se debe aplicar el proceso de ortogonalización para obtener una base seudodual derecha?
8. Demuéstrese que la descomposición de una matriz regular compleja en el producto de una matriz unitaria y una triangular derecha es única, si se exige que los elementos diagonales de la matriz triangular sean positivos.
9. ¿Cómo se emplea el proceso de ortogonalización para la resolución de los sistemas de ecuaciones algebraicas lineales?
10. Sea K_n un espacio degenerado. ¿Cómo se construirá la base seudoortogonal de un subespacio regular de dimensión máxima, empleando para ello el proceso de ortogonalización?
11. ¿Se simplificará la construcción de los sistemas seudoortogonales de vectores en un espacio degenerado, si la forma cuadrática (x, x) es de signo constante?

§ 104. Ortogonalización de una sucesión de potencias

En los procesos de ortogonalización la matriz de la transformación de coordenadas, al pasar de una base antigua a una nueva, es siempre triangular. Sin embargo, si la base inicial se elige de un modo especial, para la matriz de la transformación de coordenadas pueden obtenerse unas representaciones mucho más simples y, consecuentemente, serán más sencillos son los procesos de ortogonalización.

Supongamos que en un espacio bilineal métrico K_n, hermitiano regular, se ha dado un operador A. Tomemos un vector no nulo x y consideremos la sucesión de vectores

$$x,\ Ax,\ A^2x,\ \ldots,\ A^{n-1}x. \tag{104.1}$$

Tales sucesiones se llamarán sucesiones *de potencias* generadas por el vector x.

En toda sucesión de potencias cierto número de los primeros vectores es linealmente independiente. Supongamos que k es el mayor de estos números. Esto significa que existen tales números $\alpha_0, \alpha_1, \ldots, \alpha_k$, siendo $\alpha_k \neq 0$, que

$$\alpha_0 x + \alpha_1 Ax + \ldots + \alpha_k A^k x = 0. \tag{104.2}$$

Designemos con $\varphi(\lambda) = \alpha_k\lambda^k + \ldots + \alpha_1\lambda + \alpha_0$ el polinomio de grado k. Por lo visto, la igualdad (104.2) es equivalente a la siguiente

$$\varphi(A)\,x = 0. \tag{104.3}$$

Existen muchos polinomios, para los cuales se verifican las correlaciones del tipo (104.3). A tales polinomios pertenecen, por ejemplo, el polinomio característico del operador A. Pero entre ellos existe, a ciencia cierta, un polinomio de grado inferior. Se denomina polinomio *mínimo* que anula el vector x. Está claro que su grado es igual al número máximo de los primeros vectores de la sucesión de potencias (104.1) que forman un sistema linealmente independiente o, lo que es igual, es inferior en una unidad al número mínimo de los primeros vectores que forman un sistema linealmente dependiente.

El grado del polinomio mínimo resulta ser íntimamente relacionado con la descomposición del vector x según la base radical del operador A, las alturas de los vectores radicales y el número de los valores propios distintos dos a dos. A saber, es verídico el

Lema 104.1. *El grado del polinomio mínimo que anula el vector x es igual a la suma de las alturas máximas de los vectores radicales del operador A que figuran en la descomposición del vector x según la base radical y que corresponden a los valores propios distintos dos a dos.*

DEMOSTRACION. Representemos el vector x en forma de la suma

$$x = u_1 + u_2 + \ldots + u_s, \tag{104.4}$$

donde $u_1, \ldots, u_s$ pertenecen a los subespacios cíclicos diferentes del operador A. Puesto que los subespacios cíclicos diferentes no tienen vectores comunes, salvo el nulo, entonces para que se cumpla la igualdad (104.3), es necesario y suficiente el cumplimiento de la igualdad $\varphi(A)\,u_i = 0$ para i cualquiera. Si u_i es un vector radical de altura m_i y corresponde al valor propio λ_i, entonces la igualdad $\varphi(A)\,u_i = 0$ tendrá lugar cuando, y sólo cuando el polinomio

$\varphi(\lambda)$ se divide por $(\lambda - \lambda_i)^r$, donde $r \geqslant m_i$. En este caso $\varphi(A)\,u_j = 0$ no sólo cuando $j = i$, sino para todos los j, para los cuales los vectores u_j corresponden a los valores propios, coincidentes con λ_i, y tienen alturas no superiores a r. Sean $\lambda_{i_1}, \ldots, \lambda_{i_p}$ unos valores propios distintos dos a dos que corresponden a los vectores $u_1, \ldots, u_s$ de (104.4) y sean $m_{i_1}, \ldots, m_{i_p}$, las alturas máximas de los vectores radicales $u_1, \ldots, u_s$ que corresponden a los valores propios y coinciden con $\lambda_{i_1}, \ldots, \lambda_{i_p}$. Entonces

$$\varphi(\lambda) = (\lambda - \lambda_{i_1})^{m_{i_1}} \ldots (\lambda - \lambda_{i_p})^{m_{i_p}}$$

será el polinomio mínimo que anula el vector x. El lema queda demostrado.

Supongamos que los vectores $e_i = A^{i-1}x$, para $1 \leqslant i \leqslant k$, son linealmente independientes. Apliquemos a este sistema el proceso, descrito anteriormente, para obtener un sistema seudoortogonal de vectores f_i, considerando, por supuesto, que el propio proceso es realizable. Si el operador A de ningún modo está ligado con el producto escalar, introducido en el espacio K_n, es difícil esperar que el proceso se simplifique. Sin embargo, la situación cambia bruscamente, si el operador A satisface la correlación (101.7), siendo, por ejemplo, autoconjugado en un espacio unitario.

TEOREMA 104.1. *Si el operador A satisface la correlación* (101.7) *y si los vectores $e_i = A^{i-1}x$ son linealmente independientes para $1 \leqslant i \leqslant k$, mientras que los vectores $f_1, \ldots, f_k$ se han obtenido de los vectores $e_1, \ldots, e_k$ con ayuda del proceso de seudoortogonalización, entonces tienen lugar las siguientes correlaciones*

$$\begin{aligned} f_1 &= x, \\ f_2 &= Af_1 - \alpha_1 f_1, \\ f_{i+1} &= Af_i - \alpha_i f_i - \beta_{i-1} f_{i-1}, \quad i > 1, \end{aligned} \tag{104.5}$$

donde

$$\begin{gathered} \alpha_1 = \frac{(Af_1, f_1)}{(f_1, f_1)}, \qquad \beta_{i-1} = \frac{(Af_i, f_{i-1})}{(f_{i-1}, f_{i-1})}, \quad i > 1, \\ \alpha_i = \frac{(f_{i-1}, f_{i-1})(Af_i, f_i) - (Af_i, f_{i-1})(f_{i-1}, f_i)}{(f_{i-1}, f_{i-1})(f_i, f_i)}, \\ i > 1. \end{gathered} \tag{104.6}$$

Para concretar, la demostración se realizará en un espacio bilineal métrico hermitiano. Tomando en consideración la forma de los vectores e_i y rigiéndonos por las fórmulas (103.5), concluimos que

$$f_i = A^{i-1}x + \sum_{j=0}^{i-2} \gamma_{j,i} A^j x$$

para ciertos números $\gamma_{j,i}$. De aquí se desprende que el vector $f_{i+1} - Af_i$ pertenece a la cápsula lineal de los vectores $x, Ax, \ldots$

. . ., $A^{i-1}x$ o, lo que es igual, de los vectores $f_1, f_2, \ldots, f_i$. Por ello,

$$f_{i+1} = Af_i + \sum_{j=1}^{i} \xi_{j,\,i+1} f_j$$

para ciertos números $\xi_{j,\,i+1}$. Las condiciones de ortogonalidad del vector f_{i+1} a la izquierda respecto de los vectores $f_1, f_2, \ldots, f_i$ proporcionan el siguiente sistema de ecuaciones algebraicas lineales para determinar los coeficientes $\xi_{j,\,i+1}$

$$\begin{array}{ll}
\xi_{1,\,i+1}(f_1, f_1) & = -(Af_i, f_1), \\
\xi_{1,\,i+1}(f_1, f_2) + \xi_{2,\,i+1}(f_2, f_2) & = -(Af_i, f_2), \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
\xi_{1,\,i+1}(f_1, f_{i-1}) + \ldots + \xi_{i-1,\,i+1}(f_{i-1}, f_{i-1}) & = -(Af_i, f_{i-1}), \\
\xi_{1,\,i+1}(f_1, f_i) + \ldots + \xi_{i-1,\,i+1}(f_{i-1}, f_i) + \xi_{i,\,i+1}(f_i f_i) & = -(Af_i, f_i).
\end{array} \tag{104.7}$$

Por hipótesis del teorema, el operador A satisface la condición (101.7). Por eso, tomando en consideración la seudoortogonalidad del sistema de vectores f_j, tenemos para $l < i-1$

$$\begin{aligned}
(Af_i, f_l) &= (f_i, A^* f_l) = (f_i, (\alpha E + \beta A) f_l) = \bar{\alpha}(f_i, f_l) + \bar{\beta}(f_i, Af_l) = \\
&= \bar{\beta}(f_i, f_{l+1} - \sum_{j=1}^{l} \xi_{j,\,l+1} f_j) = \bar{\beta}\{(f_i, f_{l+1}) - \sum_{j=1}^{l} \bar{\xi}_{j,\,l+1}(f_i, f_j)\} = 0.
\end{aligned}$$

Entre los segundos miembros del sistema (104.7) solamente los dos últimos son distintos de cero. Por consiguiente, sólo $\xi_{i-1,\,i+1}$, $\xi_{i,\,i+1}$ pueden ser distintos de cero, lo que demuestra la validez de las correlaciones (104.5). El valor del coeficiente α_1 se halla, partiendo de las condiciones de ortogonalidad del vector f_2 a la izquierda respecto de f_1, mientras que los valores para los coeficientes α_i, β_{i-1} se obtienen de la condición de ortogonalidad del vector f_{i+1} a la izquierda respecto de f_i, f_{i-1}.

Así pues, el proceso de ortogonalización para una sucesión de potencias resulta realmente mucho más simple que para una sucesión de forma general. Si en cierto paso resulta que $(f_i, f_i) = 0$, pero $f_i \neq 0$, entonces la degeneración del proceso puede ser evitada por elección adecuada del vector nuevo x.

Supongamos que en una sucesión de potencias se tienen n vectores linealmente independientes. Aplicando el proceso de ortogonalización, se puede construir una base $f_1, f_2, \ldots, f_n$ del espacio K_n. Esta base posee la singularidad de que en ella la matriz A_f del operador A

tiene forma *tridiagonal*

$$A_f=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ 1 & \alpha_2 & \beta_2 & & \\ & 1 & \alpha_3 & \beta_3 & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & & 1 & \alpha_{n-1} & \beta_{n-1} \\ & & & 1 & \alpha_n \end{pmatrix}. \qquad (104.8)$$

En efecto, como columnas de la matriz del operador actúan las coordenadas de los vectores $Af_1, Af_2, \ldots, Af_n$ respecto de la base $f_1, f_2, \ldots$ $\ldots, f_n$. Pero, de acuerdo con (104.5),

$$\begin{aligned} Af_1 &= \alpha_1 f_1 + f_2, \\ Af_2 &= \beta_1 f_1 + \alpha_2 f_2 + f_3, \\ &\cdots\cdots\cdots\cdots\cdots \\ Af_{n-1} &= \beta_{n-2} f_{n-2} + \alpha_{n-1} f_{n-1} + f_n, \\ Af_n &= \beta_{n-1} f_{n-1} + \alpha_n f_n. \end{aligned} \qquad (104.8)$$

Si en una sucesión de potencias faltan n vectores linealmente independientes, entonces, aplicando el proceso de ortogonalización, llegamos a que $f_{r+1}=0$ para cierto $r<n$. Tomemos un vector nuevo u y formemos el vector

$$v=u-\sum_{j=1}^{r}\eta_j f_j,$$

determinando los coeficientes η_j de las condiciones de ortogonalidad del vector v a la izquierda respecto de los vectores $f_1, f_2, \ldots, f_r$. Construyamos una sucesión de potencias engendrada por el vector v. Es fácil probar que todo vector de esta sucesión es también ortogonal a la izquierda respecto de los vectores $f_1, f_2, \ldots, f_r$. Dicha propiedad la posee por construcción el vector v. Supongamos que la hemos demostrado para todos los vectores $v, Av, \ldots, A^k v$. En este caso, teniendo presentes las correlaciones (104.9) y la igualdad (101.7), obtenemos para el operador A

$$\begin{aligned}(A^{k+1}v, f_l) &= (A^k v, A^* f_l) = (A^k v, (\alpha E+\beta A) f_l) = \\ &= \bar{\alpha}(A^k v, f_l) + \bar{\beta}(A^k v, \beta_{l-1} f_{l-1} + \alpha_l f_l + f_{l+1}) = 0.\end{aligned}$$

Aplicando el proceso de ortogonalización a la nueva sucesión, construiremos un sistema de vectores linealmente independientes $q_1, q_2, \ldots, q_s$ que son seudoortogonales entre sí y a la totalidad de los vectores $f_1, f_2, \ldots, f_r$. Si $r+s<n$, continuaremos el proceso de construcción de la base hasta que se haya construido la base de todo el espacio. En este caso todo el espacio se dividirá en la suma directa

de los subespacios invariantes, y la matriz del operador A en dicha base será de forma celular diagonal con células tridiagonales del tipo (104.8).

Ejercicios.

1. Demuéstrese que el polinomio mínimo que anula el vector x es un divisor del polinomio característico.

2. Demuéstrese que el polinomio mínimo que anula el vector x es único, salvo un factor escalar.

3. Demuéstrese que si el operador A es hermitiano y el espacio K_n, unitario, entonces las fórmulas (104.6) adquieren la forma:

$$\alpha_i = \frac{(Af_i, f_i)}{(f_i, f_i)},$$

$$\beta_{i-1} = \frac{(Af_i, f_{i-1})}{(f_{i-1}, f_{i-1})} = \frac{(f_i, Af_{i-1})}{(f_{i-1}, f_{i-1})} = \frac{(f_i, f_i)}{(f_{i-1}, f_{i-1})} > 0.$$

4. Demuéstrese que en las condiciones del ejercicio 3 existe tal matriz diagonal D que para la matriz A_f de (104.8) la matriz $D^{-1}A_fD$ será real simétrica tridiagonal.

5. Demuéstrese que, cumplidas las condiciones (101 7), las matrices de las formas bilineales (Ax, y), (x, Ay) en la base $f_1, \ldots, f_n$ son derechas casi triangulares.

6. Demuéstrese que en las condiciones del ejercicio 3 las matrices de las formas bilineales (Ax, y), (x, Ay) en la base $f_1, \ldots, f_n$ son hermitianas tridiagonales.

7. Demuéstrese que si el espacio K_n es degenerado, entonces, con ayuda de los procesos (104.5), (104.6), puede construirse la base seudoortogonal de un subespacio regular de dimensión máxima.

105. Métodos de direcciones conjugadas

La construcción de los sistemas de vectores ortogonales, seudoortogonales y otros, especialmente a la base del empleo de las sucesiones de potencias, ofrece grandes posibilidades en la elaboración de toda una serie de métodos numéricos para solucionar las ecuaciones del tipo

$$Ax = b, \tag{105.1}$$

donde A es un operador que actúa en el espacio lineal K_n, mientras que b es un vector prefijado y x, el vector buscado.

Ya nos hemos referido reiteradamente a los diferentes aspectos de este problema. Ahora daremos a conocer un gran grupo de métodos numéricos para la resolución de la ecuación (105.1) a los que se ha atribuido el nombre general de métodos de direcciones conjugadas. Todos ellos están basados en los procesos de ortogonalización de sucesiones de potencias. Con el fin de simplificar la exposición, supongamos que el operador A es regular y, por lo tanto, la ecuación

(105.1) tiene siempre la única solución. Convengamos en considerar que el espacio K_n es complejo y el producto escalar en él viene dado mediante una forma bilineal hermitiana simétrica y definida positiva, es decir, K_n es un espacio unitario.

Tomemos unos operadores regulares C, B cualesquiera y sea $s_1, \ldots, s_n$ un sistema de vectores CAB-seudoortogonal, es decir

$$(CABs_i, s_i) \neq 0, \qquad (CABs_i, s_k) = 0, \quad k < i,$$

para todo i. Designemos con x_0 el vector inicial y sea

$$x = x_0 + B \sum_{j=1}^{n} a_j s_j,$$

$$x_i = x_0 + B \sum_{j=1}^{i} a_j s_j, \tag{105.2}$$

$$r_i = Ax_i - b.$$

Entonces, de las correlaciones

$$x_i = x_{i-1} + a_i B s_i \tag{105.3}$$

se infiere que

$$r_i = r_{i-1} + a_i A B s_i. \tag{105.4}$$

Es fácil mostrar que para el sistema CAB-seudoortogonal $s_1, \ldots$ $\ldots, s_n$ tienen lugar las igualdades

$$(Cr_i, s_k) = 0, \qquad 1 \leqslant k \leqslant i. \tag{105.5}$$

En efecto,

$$r_i = Ax_i - b = A(x_i - x) = -\sum_{j=i+1}^{n} a_j A B s_j$$

y, además,

$$(Cr_i, s_k) = -\sum_{j=i+1}^{n} a_j (CABs_j, s_k) = 0$$

para cualesquiera $k \leqslant i$.

Consideraremos que el sistema de vectores $s_1, \ldots, s_n$ se construye paralelamente con el sistema $r_0, \ldots, r_{n-1}$, con ayuda del proceso de su CAB-seudoortogonalización. Hagamos $s_1 = r_0$ y para todo i tendremos

$$s_{i+1} = r_i + \sum_{k=1}^{i} \beta_{k,\, i+1} s_k. \tag{105.6}$$

Las condiciones de CAB-ortogonalidad del vector s_{i+1} a la izquierda respecto de los vectores $s_1, \ldots, s_i$ proporcionan, como siempre, un sistema triangular izquierdo para determinar los coeficientes $\beta_{k,i+1}$. En este caso, r_k es una combinación lineal de los vectores $s_1, \ldots$ $\ldots, s_{k+1}$. Por consiguiente, el producto escalar (Cr_i, r_k) es una combinación lineal de los números $(Cr_i, s_1), \ldots, (Cr_i, s_{k+1})$ y es igual a cero, de conformidad con (105.5) para $k < i$, es decir,

$$(Cr_i, r_k) = 0, \qquad k < i. \tag{105.7}$$

Esto significa que si $(Cr_i, r_i) \neq 0$ para i cualquiera, entonces la sucesión de vectores r_i es C-seudoortogonal. En un espacio lineal n-dimensional un sistema C-seudoortogonal no puede contener más de n vectores no nulos. Por esta razón, en cierto paso del proceso de cálculo uno de los residuos se hará nulo y obtendremos la solución exacta de la ecuación (105.1).

Para realizar el proceso, es menester determinar los coeficientes a_i de (105.2) y $\beta_{k,i+1}$ de (105.6). Los coeficientes a_i se hallan siempre de un modo simple. Conforme a (105.4), (105.5), (105.7), tenemos

$$a_i = -\frac{(Cr_{i-1}r_{i-1})}{(CABs_i, r_{i-1})} = -\frac{(Cr_{i-1}, s_i)}{(CABs_i, s_i)}. \tag{105.8}$$

En el caso general, los coeficientes $\beta_{k,i+1}$ se calculan de una manera mucho más compleja. Sin embargo, si los operadores A, B, C están ligados entre sí por la correlación

$$(CABC^{-1})^* = \alpha E + \beta AB \tag{105.9}$$

para ciertos números α, β, entonces entre todos los coeficientes $\beta_{k,i+1}$ sólo $\beta_{i,i+1}$ puede ser distinto de cero. Supongamos que

$$s_{i+1} = r_i + b_i s_i. \tag{105.10}$$

El coeficiente b_i se determina unívocamente de la condición de CAB-ortogonalidad del vector s_{i+1} a la izquierda respecto del vector s_i, lo que nos da, al tomar en consideración (105.9), (105.10)

$$b_i = -\frac{(CABr_i, s_i)}{(CABs_i, s_i)} = -\bar{\beta}\,\frac{(Cr_i, ABs_i)}{(CABs_i, s_i)}. \tag{105.11}$$

Supongamos que al calcular la sucesión de los vectores s_i según las fórmulas (105.10), (105.11), hemos mostrado que la sucesión $s_1, \ldots, s_i$ forma un sistema CAB-seudoortogonal. Esto es justo, a ciencia cierta, cuando $i = 2$. Teniendo presente (105.9), obtenemos de (105.4)—(105.7), para $k < i$, que

$$\begin{aligned}(CABs_{i+1}, s_k) &= (CABr_i, s_k) + b_i\,(CABs_i, s_k) = \\ &= ((CABC^{-1})\,Cr_i s_k) = (Cr_i, (CABC^{-1})^* s_k) = \\ &= (Cr_i, (\alpha E + \beta AB)\,s_k) = \bar{\alpha}\,(Cr_i, s_k) + \bar{\beta}\,(Cr_i, ABs_k) = \\ &= \bar{\beta}\left(Cr_i, \frac{1}{a_k}(r_k - r_{k-1})\right) = \frac{\bar{\beta}}{\bar{a}_k}\{(Cr_i, r_k) - (Cr_i, r_{k-1})\} = 0.\end{aligned}$$

De este modo, si se cumple la correlación (105.9), la resolución de la ecuación operacional (105.1) puede realizarse a base de la siguiente sugestión:

$$\begin{aligned} s_1 &= r_0, \\ r_i &= r_{i-1} + a_i ABs_i, \\ s_{i+1} &= r_i + b_i s_i, \\ x_i &= x_{i-1} + a_i Bs_i. \end{aligned} \tag{105.12}$$

Aquí, x_0 es un vector inicial arbitrario y los coeficientes a_i, b_i se calculan de acuerdo con (105.8), (105.11). Al designar $u_i = Bs_i$, el proceso será:

$$\begin{aligned} u_1 &= Br_0, \\ r_i &= r_{i-1} + a_i Au_i, \\ u_{i+1} &= Br_i + b_i u_i, \\ x_i &= x_{i-1} + a_i u_i \end{aligned}$$

siendo para este caso

$$a_i = -\frac{(Cr_{i-1}, r_{i-1})}{(CAu_i, r_{i-1})} = -\frac{(B^{-1*}Cr_{i-1}, u_i)}{(B^{-1*}CAu_i, u_i)},$$

$$b_i = -\frac{(B^{-*1}CABr_i, u_i)}{(B^{-*1}CAu_i, u_i)} = -\bar{\beta}\frac{(Cr_i, Au_i)}{(B^{-1*}CAu_i, u_i)}.$$

Estos procesos llevan el nombre de *métodos de direcciones conjugadas.*

De las fórmulas (105.4), (105.10) concluimos que los vectores r_i, s_{i+1} son combinaciones lineales de los vectores de una misma sucesión de potencias

$$r_0, \ ABr_0, \ \ldots, \ (AB)^i r_0. \tag{105.13}$$

Más aún, se han obtenido de ésta con ayuda de C- y CAB-seudoortogonalización, respectivamente. De este hecho se deducen unos corolarios de importancia excluvisa.

Si en la descomposición del vector r_0 según la base canónica de Jordan del operador AB no todos los componentes están presentes, entonces la anulación del residuo sucederá antes que llegue el n-ésimo paso. El proceso se termina con una rapidez singular, si el operador AB es de estructura simple y tiene un gran número de valores propios coincidentes. A saber, si en la descomposición del vector r_0 según los valores propios de la matriz AB los componentes no nulos corresponden a m valores propios distintos dos a dos, entonces $r_m = 0$.

De acuerdo con el teorema 104.1, para los vectores s_i, r_i deben tener lugar las correlaciones de tres términos del tipo (104.5). Pueden obtenerse éstas directamente de (105.4), (105.10). A saber,

$$\begin{aligned} s_{i+1} &= a_i ABs_i + (1 + b_i) s_i - b_{i-1} s_{i-1}, \quad i > 1, \\ r_{i+1} &= a_{i+1} ABr_i + \left(1 + \frac{b_i a_{i+1}}{a_i}\right) r_i - \frac{b_i a_{i+1}}{a_i} r_{i-1} \quad i \geqslant 1. \end{aligned} \tag{105.14}$$

De aquí pueden obtenerse también otras correlaciones. Por ejemplo,

$$x_{i+1} = x_{i-1} + \omega_{i+1} (\alpha_i B r_i + x_i - x_{i-1}),$$

donde ω_{i+1}, α_i son unos números elegidos de modo adecuado.

En relación con lo dicho respecto de la sucesión (105.13), fijemos nuestra atención en la siguiente peculiaridad de la condición (105.9). A primera vista, se diferencia de las condiciones del tipo (101.7). Sin embargo, al tomar en consideración (101.6), resulta fácil mostrar que la condición (105.9) es, de hecho, también una condición del tipo (101.7) y, además, simultáneamente respecto a dos productos escalares $(CABx, y)$ y (Cx, y). En efecto, observemos que el operador conjugado en (105.9) está ligado con el próducto escalar principal del espacio unitario, mientras que la ortogonalidad de los vectores s_i, r_i se asegura respecto de los productos escalares $(CABx, y)$ y (Cx, y), respectivamente. Por ello

$${}_{CAB}(AB)^* = (CAB \cdot AB \cdot (CAB)^{-1})^* = (CABC^{-1})^* = \alpha E + \beta AB,$$

$${}_{C}(AB)^* = (CABC^{-1})^* = \alpha E + \beta AB.$$

La realización de los métodos de direcciones conjugadas puede ser obstacularizada sólo por el hecho de que uno de los productos escalares, $(CABs_i, s_i)$ o (Cr_{i-1}, r_{i-1}), se anule antes de que se reduzca a cero el residuo. Si $(CABs_i, s_i) = 0$, los coeficientes a_i, b_i no pueden calcularse. En cambio, si $(Cr_{i-1}, r_{i-1}) = 0$, esto tendrá por resultado que el coeficiente a_i será nulo, mientras que los residuos no nulos r_{i-1}, r_i coincidirán y, como consecuencia, tendrá lugar la igualdad $(CABs_{i+1}, s_{i+1}) = 0$. Se puede evitar esta situación mediante la elección del nuevo vector inicial x_0. Si los operadores CAB y C son definidos positivos, las degeneraciones citadas son imposibles y el proceso fluye sin complicaciones. Si el operador CAB es definido positivo, los métodos de direcciones conjugadas adquieren adicionalmente unas nuevas propiedades interesantes.

Sobre el grado en que el vector z se aproxima a la solución de la ecuación (105.1) podemos juzgar por la pequeñez del cuadrado de una norma de la diferencia $e = x - z$. Con este fin resulta cómodo utilizar las así llamadas *funcionales generalizadas de errores* del tipo (Re, e), donde R es cualquier operador definido positivo, por ejemplo, el operador $B^{-1*}CA$. El operador dado será definido positivo, puesto que está asociado al operador CAB mediante la relación $B^{-1*}CA = B^{-1*} (CAB) B^{-1}$. Es válido el

TEOREMA 105.1. *Si el operador CAB es definido positivo, entre todos los vectores del tipo $z = x_0 + Bs$, donde s pertenece a la cápsula lineal de los vectores $s_1, \ldots, s_i$, el vector x_i da el mínimo de la funcional generalizada de errores*

$$\varphi(z) = (B^{-1*}CAe, e).$$

DEMOSTRACION. Puesto que el operador CAB es definido positivo, el sistema de los vectores s_i será CAB-ortogonal. Representemos el vector z en forma de una descomposición, por analogía con la descomposición (105.2) para el vector x:

$$z = x_0 + B\sum_{j=1}^{i} h_j s_j.$$

Tenemos

$$\varphi(z) = (B^{-1*}CA(x-z),\ x-z) =$$
$$= \Big(B^{-1*}CAB\big(\sum_{j=1}^{n} a_j s_j - \sum_{j=1}^{i} h_j s_j\big),\ B\big(\sum_{j=1}^{n} a_j s_j - \sum_{j=1}^{i} h_j s_j\big)\Big) =$$
$$= \Big(CAB\big(\sum_{j=1}^{n} a_j s_j - \sum_{j=1}^{i} h_j s_j\big),\ \sum_{j=1}^{n} a_j s_j - \sum_{j=1}^{i} h_j s_j\Big) =$$
$$= \sum_{j=1}^{i} |a_j - h_j|^2 (CABs_j,\ s_j) + \sum_{j=i+1}^{n} |a_j|^2 (CABs_j,\ s_j).$$

De aquí concluimos que el mínimo de la funcional de errores se alcanza para $h_j = a_j$, $j \leqslant i$, es decir, para $z = x_i$.

La funcional de errores no puede determinarse en los cálculos prácticos, puesto que depende de la solución x que se desconoce. No obstante, dicha funcional sólo se diferencia de otra funcional en un umando constante:

$$\psi(z) = (B^{-1*}CAz,\ z) - 2\mathrm{Re}\,(B^{-1*}Cb,\ z)$$

que ya puede calcularse. En efecto,

$$\varphi(z) = (B^{-1*}CA(x-z),\ x-z) =$$
$$= (B^{-1*}CAx,\ x) - (B^{-1*}CAx,\ z) - (B^{-1*}CAz,\ x) +$$
$$+ (B^{-1*}CAz,\ z) = (B^{-1*}CAz,\ z) - (B^{-1*}Cb,\ z) -$$
$$- \overline{(B^{-1*}Cb,\ z)} + (B^{-1*}Cb,\ x) = \psi(z) + (B^{-1*}Cb,\ x).$$

Demos a conocer, en fin, algunas clases de los operadores, para los cuales se cumplen las condiciones (105.9).

1. Todos los operadores A, B, C son hermitianos, siendo $B = C$. La condición (105.9) se cumple para $\alpha = 0$, $\beta = 1$:

$$(CABC^{-1})^* = (BA)^* = A^*B^* = 0\cdot E + 1\cdot AB.$$

2. Los operadores CAB y C son hermitianos. La condición (105.9) se cumple para $\alpha = 0$, $\beta = 1$:

$$(CABC^{-1})^* = C^{-1*}(CAB)^* = C^{-1}CAB = 0\cdot F + 1\cdot AB.$$

3. El operador C se conmuta con AB; el operador AB es normal y su espectro se halla en una línea recta. La última condición signi-

fica que $AB = \gamma E + \delta H$ para cierto operador hermitiano H. Ahora hallamos

$$(CABC^{-1})^* = (CC^{-1}AB)^* = (AB)^* = (\gamma E + \delta H)^* =$$
$$= \bar{\gamma} E + \bar{\delta} H = \frac{\bar{\gamma}\delta - \bar{\delta}\gamma}{\delta} E + \frac{\bar{\delta}}{\delta}(\gamma E + \delta H) = + \frac{2\,\mathrm{Im}\,(\bar{\gamma}\delta)}{\delta} E + \frac{\bar{\delta}}{\delta} AB.$$

4. Representemos el operador A en la forma $A = M + N$, donde $M = M^*$, $N = -N^*$. Si el operador M es regular, hacemos $B = C = M^{-1}$. La condición (105.9) se cumple cuando $\alpha = 2$, $\beta = -1$:

$$(CABC^{-1})^* = (M^{-1}(M + N))^* = (M - N) M^{-1} =$$
$$= 2E - (M + N) M^{-1} = 2 \cdot E - 1 \cdot AB.$$

5. Si el operador N en la descomposición $A = M + N$ es regular, hacemos $B = C = N^{-1}$. La condición (105.9) se cumple para $\alpha = 2$, $\beta = -1$:

$$(CABC^{-1})^* = (N^{-1}(M + N))^* = -(M - N) N^{-1} =$$
$$= 2E - (M + N) N^{-1} = 2 \cdot E - 1 \cdot AB.$$

Ejercicios.

1. Demuéstrese que la matriz de la forma bilineal $(CABx, y)$ es:
triangular derecha en la base $s_1, \ldots, s_n$,
casi triangular derecha en la base $r_0, \ldots, r_{n-1}$,
triangular izquierda en las bases $s_1, \ldots, s_n$ y $r_0, \ldots, r_{n-1}$, si el operador C es hermitiano.

2. ¿Cómo varía la forma de una matriz de la forma bilineal $(CABx, y)$ en el ejercicio 1, si el operador CAB es hermitiano?

3. Demuéstrese que la matriz de la forma bilineal (Cx, y) es:
triangular derecha en la base $r_0, \ldots, r_{n-1}$,
triangular derecha en las bases $r_0, \ldots, r_{n-1}$ y $s_1, \ldots, s_n$,
triangular derecha en las bases $ABs_1, \ldots, ABs_n$ y $s_1, \ldots, s_n$,
casi triangular derecha en las bases $ABr_0, \ldots, ABr_{n-1}$ y $r_0, \ldots, r_{n-1}$.

4. ¿Cómo varía la forma de una matriz de la forma bilineal (Cx, y) en los ejercicios 3, si el operador C es hermitiano?

5. Demuéstrese que si la condición (105.9) se sustituye por otra:

$$(CABC^{-1})^* = \alpha_0 E + \alpha_1 AB + \ldots + \alpha_p (AB)^p, \qquad p \geqslant 1,$$

la correlación (105.10) tendrá por expresión

$$s_{i+1} = r_i + b_i s_i + b_{i-1} s_{i-1} + \ldots + b_{i-p+1} s_{i-p+1}.$$

6. Demuéstrese que

$$a_i = -\frac{(Cr_{i-1}, r_{i-1})}{(CABs_i, s_i)}.$$

7. Demuéstrese que si los operadores CAB y C son hermitianos, se verifica

$$b_i = \frac{(Cr_i, r_i)}{(Cr_{i-1}, r_{i-1})}.$$

8. Demuéstrese que si los operadores CAB y C son hermitianos y, además, definidos positivos, entonces $a_i < 0$, $b_i > 0$ para todo valor de i.

9. Demuéstrese que la matriz del operador AB tiene en una base de los vectores $s_1, \ldots, s_n$ ó $r_0, \ldots, r_{n-1}$ la forma tridiagonal.
10. ¿Cómo se realizan los métodos de direcciones conjugadas en el caso cuando el operador A es degenerado?

§ 106. Variantes principales

Consideraremos las variantes más conocidas de los métodos de direcciones conjugadas. En el aspecto teórico todas ellas se encajan dentro del esquema descrito anteriormente (105.12). Sin embargo, los cálculos prácticos se efectúan, a veces, a base de los algoritmos algo diferentes de éste.

Método de gradientes conjujados. En este caso el operador A es hermitiano y, además, definido positivo; $B = C = E$; la condición (105.9) se cumple para $\alpha = 0$, $\beta = 1$. Debido a que los operadores $CAB = A$ y $C = E$ son definidos positivos se garantiza la ausencia de degeneraciones en el proceso de cálculo. En cada etapa del método se minimiza la funcional de errores provista de la matriz A. El esquema de cálculo de este método es como sigue

$$\begin{aligned} s_1 &= r_0, \\ r_i &= r_{i-1} + a_i A s_i, \\ s_{i+1} &= r_i + b_i s_i, \\ x_i &= x_{i-1} + a_i s_i, \end{aligned}$$

donde

$$a_i = -\frac{(r_{i-1}, r_{i-1})}{(As_i, r_{i-1})} = -\frac{(r_{i-1}, s_i)}{(As_i, s_i)} < 0,$$

$$b_i = -\frac{(r_i, As_i)}{(As_i, s_i)} = \frac{(r_i, r_i)}{(r_{i-1}, r_{i-1})} > 0.$$

Al aplicar el método de gradientes conjugados los vectores r_i forman un sistema ortogonal y los vectores s_i, un sistema A-ortogonal.

Método de iteraciones AA^*-minimales. En este caso el operador A es regular arbitrario; $B = A^*$, $C = E$; la condición (105.9) se cumple para $\alpha = 0$, $\beta = 1$. El hecho de que los operadores $CAB = AA^*$ y $C = E$ son definidos positivos garantiza la ausencia de las degeneraciones en el proceso de cálculo. En cada etapa del método se minimiza la funcional de errores con la matriz A, es decir, el cuadrado de la norma euclidiana del mismo error. El esquema de cálculo es como sigue

$$\begin{aligned} u_1 &= A^* r_0, \\ r_i &= r_{i-1} + a_i A u_i, \\ u_{i+1} &= A^* r_i + b_i u_i, \\ x_i &= x_{i-1} + a_i u_i, \end{aligned}$$

donde

$$a_i = -\frac{(r_{i-1}, r_{i-1})}{(Au_i, r_{i-1})} = -\frac{(r_{i-1}, r_{i-1})}{(u_i, u_i)} < 0,$$

$$b_i = -\frac{(r_i, Au_i)}{(u_i, u_i)} = \frac{(r_i, r_i)}{(r_{i-1}, r_{i-1})} < 0.$$

Al aplicar el método de iteraciones AA^*-minimales los vectores r_i y u_i forman sistemas ortogonales.

Método de iteraciones A^*A-minimales. En este caso el operador A es regular arbitrario; $B = A^*$, $C = AA^*$; la condición (105.9) se cumple para $\alpha = 0$, $\beta = 1$. El hecho de que los operadores $CAB = (AA^*)^2$ y $C = AA^*$ son definidos positivos garantiza la ausencia de las degeneraciones en el proceso de cálculo. En cada etapa del método se minimiza la funcional de errores provista de la matriz A^*A, es decir, el cuadrado de la norma euclidiana del vector del residuo. El esquema de cálculo es como sigue

$$u_1 = A^*r_0,$$
$$r_i = r_{i-1} + a_iAu_i,$$
$$u_{i+1} = A^*r_i + b_iu_i,$$
$$x_i = x_{i-1} + a_iu_i,$$

de donde

$$a_i = -\frac{(A^*r_{i-1}, A^*r_{i-1})}{(Au_i, Au_i)} < 0, \qquad b_i = \frac{(A^*r_i, A^*r_i)}{(A^*r_{i-1}, A^*r_{i-1})} > 0.$$

Al aplicar el método de iteraciones A^*A-minimales los vectores A^*r_i y Au_i forman sistemas ortogonales.

Método de descomposición hermitiana completa. En este caso el operador A es regular arbitrario. Representémoslo en forma de una suma $A = M + N$, donde $M = M^*$, $N = -N^*$. En el caso en que M o N sea regular, ponemos $B = C = M^{-1}$ o $B = C = N^{-1}$, respectivamente. La condición (105.9) se cumple para $\alpha = 2$, $\beta = -1$. Si el operador M (o iN) es de signo constante, el proceso se realiza sin degeneraciones. Sea, por ejemplo, $M > 0$ y $B = C = M^{-1}$. El operador C será definido positivo, por lo tanto $(Cz, z) = (M^{-1}, z, z) > 0$, para cualquier vector z no nulo. Consideremos ahora el operador $CAB = M^{-1} + M^{-1}NM^{-1}$. Para cualquier $z \neq 0$ tenemos

$$(CABz, z) = (M^{-1}z, z) + (M^{-1}NM^{-1}z, z) \neq 0,$$

puesto que el primer producto escalar del segundo miembro de la igualdad es real y positivo, por ser definido positivo el operador, mientras que el segundo producto escalar es imaginario puro, en virtud de que el operador $M^{-1}NM^{-1}$ es antihermitiano. Para el caso en que $B = C = M^{-1}$, el esquema de cálculo del método es como sigue

$$\begin{aligned} Mu_1 &= r_0, \\ r_i &= r_{i-1} + a_i A u_i, \\ Mv_i &= r_i, \\ u_{i+1} &= v_i + b_i u_i, \\ x_i &= x_{i-1} + a_i u_i, \end{aligned}$$

donde

$$a_i = -\frac{(r_{i-1}, u_i)}{(Au_i, u_i)}, \qquad b_i = \frac{(v_i, Au_i)}{(Au_i, u_i)}.$$

Si $B = C = N^{-1}$, el esquema de cálculo y las fórmulas para los coeficientes a_i, b_i quedan en vigor, a excepción de que M se sustituye por N.

Método de descomposición hermitiana incompleta. En este caso el operador A es hermitiano, definido positivo. Representémoslo en forma de una suma $A = M + N$, donde $M = M^*$, $N = N^*$. Si M es regular, ponemos $B = C = M^{-1}$. La condición (105.9) se cumple cuando $\alpha = 0$, $\beta = 1$. Si M es un operador definido positivo, el proceso se realiza sin degeneraciones. En cada etapa del método se minimiza la funcional de errores provista de la matriz A. El esquema de cálculo queda el mismo que en el caso antecedente.

Aceleración del proceso de cálculo. Como ya se ha notado anteriormente, los métodos de direcciones conjugadas permiten hallar la solución con una rapidez singular, si el operador AB cuenta con pocos valores propios distintos dos a dos. Esta singularidad sirve de base para diferentes procedimientos que tienen por objeto acelerar la resolución de la ecuación (105.1), empleando la siguiente idea.

Supongamos que el operador A puede representarse en forma de una suma $A = M + N$, donde el operador M determina la parte «principal» del operador A y admite, además, una resolución sencilla de las ecuaciones del tipo (105.1) con el operador M en el primer miembro. Ahora, en lugar de la ecuación (105.1) resolveremos la ecuación

$$(E + NM^{-1})\, y = b, \tag{106.1}$$

donde $Mx = y$. Si, en algún sentido razonable, el operador M es próximo al operador A, entonces la mayoría de los valores propios del operador N y, por tanto, del operador NM^{-1} serán próximos a cero o iguales a cero. La aplicación de los métodos de direcciones conjugadas a la ecuación (106.1) conduce, en el caso dado, a su rápida resolución.

Ha de ser notado que precisamente esta idea es la base para crear un método de descomposición hermitiana incompleta, el cual en muchos casos resulta más eficaz que el de gradientes en la variante clásica. Todo depende de cuán exitosa ha sido la descomposición del operador A.

No es nuestra intención detenernos detalladamente en los esquemas de cálculo para los procesos de aceleración, pues dependen demasiado del empleo de unas u otras peculiaridades del operador A.

Ejercicios.

1. ¿En qué condiciones resulta conveniente aplicar una u otra variante del método de direcciones conjugadas?

2. ¿Qué número de iteraciones se deben cumplir, realizando diversas variantes de los métodos de direcciones conjugadas para un operador A del tipo $E + R$, donde R es de rango r?

3. Considerando el operador A una matriz, evalúese el número de operaciones aritméticas que se deben cumplir al resolver sistemas de las ecuaciones algebraicas lineales por los métodos de direcciones conjugadas.

4. La matriz del operador A es hermitiana y se diferencia de la matriz triangular por su número pequeño de elementos. ¿Cuál de las variantes de los métodos de direcciones conjugadas es la más conveniente para el empleo en el caso dado?

5. Sea $P_0(t)$, $P_1(t)$, . . . una sucesión de polinomios. Elijamos un vector x_0 y construyamos una sucesión de vectores x_0, x_1, . . ., rigiéndonos por la regla

$$x_{k+1} = x_k - BP_k(AB)(Ax_k - b), \qquad k \geqslant 0. \tag{106.2}$$

¿Cómo varían las descomposiciones de los residuos r_0, r_1, . . . según la base canónica de Jordan del operador AB, cuando k crece en función de la elección de la sucesión de polinomios?

6. ¿Cómo debe utilizarse la sucesión (106.2) con el fin de construir un vector inicial para los métodos de direcciones conjugadas el cual asegure la obtención de la solución por una cantidad menor de iteraciones?

7. ¿Cuáles de los sistemas de vectores en cada una de las variantes concretas de los métodos de direcciones conjugadas son, salvo una normalización, A-seudodual?

§ 107. Ecuaciones operacionales y seudodualidad

Los métodos de direcciones conjugadas no son únicos entre aquellos que se emplean para la resolución de la ecuación operacional

$$Ax = b, \tag{107.1}$$

y están basados en la aplicación de las formas bilineales. Enormes posibilidades para crear estos métodos ofrece la construcción de los sistemas de vectores, duales o seudoduales respecto a cierta forma bilineal, vinculada con el operador A de la ecuación (107.1).

Volvamos a considerar que el operador A es regular y actúa en un espacio unitario K_n. Examinemos una forma bilineal (Ax, y) y supongamos que para ésta se han obtenido, de una u otra manera, los sistemas de vectores $u_1, u_2, \ldots, u_n$ y $v_1, v_2, \ldots, v_n$ que son A-seudoduales, salvo una normalización, es decir,

$$(Au_i, v_i) \neq 0, \qquad (Au_i, v_k) = 0, \quad k < i. \tag{107.2}$$

para todo valor de $i \leqslant n$. Probemos que el conocimiento de los sistemas A-seudoduales de vectores permite construir un proceso de búsqueda de la solución de la ecuación (107.1).

Elijamos un vector arbitrario x_0. Como los sistemas A-seudoduales son linealmente independientes, existe la descomposición

$$x = x_0 + \sum_{j=1}^{n} a_j u_j. \tag{107.3}$$

Si

$$x_i = x_0 + \sum_{j=1}^{i} a_j u_j,$$

entonces, por analogía con (105.3), (105.4), tenemos

$$x_i = x_{i-1} + a_i u_i, \qquad r_i = r_{i-1} + a_i A u_i. \tag{107.4}$$

Luego,

$$r_i = Ax_i - b = A(x_i - x) = -\sum_{j=i+1}^{n} a_j A u_j,$$

y, en plena concordancia con las segundas condiciones (107.2), encontramos que

$$(r_i, v_k) = -\sum_{j=i+1}^{n} a_i (Au_j, v_k) = 0$$

para todo valor de $k \leqslant i$. Así pues,

$$(r_i, v_k) = 0 \tag{107.5}$$

para $k \leqslant i$. Esto nos permite determinar los coeficientes a_i de (107.4), a saber

$$a_i = \frac{(r_{i-1}, v_i)}{(Au_i, v_i)}. \tag{107.6}$$

De acuerdo con la primera condición (107.2), el denominador en el segundo miembro de (107.6) es distinto de cero.

De (107.5) se deduce que el vector r_n será ortogonal a la izquierda y, por ser simétrico el producto escalar, simplemente ortogonal a los vectores linealmente independientes $v_1, v_2, \ldots, v_n$ es decir, $r_n = 0$, mientras que x_n es la solución de la ecuación (107.1).

Los métodos descritos de resolución de la ecuación (107.1) llevan el nombre general de métodos de direcciones duales. El número de diferentes métodos es infinito en pleno sentido de esta palabra, puesto que existe un número infinito de distintos pares A-seudoduales de los sistemas de vectores. Los métodos de direcciones conjugadas, considerados anteriormente, forman parte de este grupo.

En el caso general, para los métodos de direcciones duales no existe ningún análogo del teorema 105.1, incluso cuando el operador A sea definido positivo. La ley de variación de los errores $e_k = x - x_k$ en estos métodos describe el teorema siguiente que, aunque débil, es sin embargo útil.

TEOREMA 107.1 *Sea P_k un operador de proyección sobre el subespacio tendido sobre los vectores $u_1, \ldots, u_k$ paralelamente a un subespacio tendido sobre los vectores $u_{k+1}, \ldots, u_n$. Entonces*

$$e_k = (E - P_k)\, e_0. \tag{107.7}$$

DEMOSTRACION. De acuerdo con la fórmula (107.3), tenemos la siguiente descomposición del vector inicial x_0 para el error e_0:

$$e_0 = x - x_0 = \sum_{j=1}^{n} a_j u_j.$$

Pero, por definición del operador de proyección,

$$P_k e_0 = \sum_{j=1}^{k} a_j u_j.$$

El segundo miembro de esta igualdad no es otra cosa que $x_k - x_0$. Por esto

$$P_k e_0 = x_k - x_0 = (x - x_0) - (x - x_k) = e_0 - e_k,$$

lo que demuestra la afirmación del teorema.

Unos resultados interesantes relacionados con los sistemas A-seudoduales pueden obtenerse, considerando la interpretación matricial de los métodos descritos.

Consideraremos que el espacio K_n no es sólo unitario, sino también aritmético, que es admisible en virtud de que los espacios lineales de dimensión finita son isomorfos. Todos los razonamientos realizados quedan en vigor, sólo cambia la terminología: la ecuación (101.7) pasa a ser un sistema de ecuaciones algebraicas lineales, los operadores se sustituyen por matrices y por vectores se entienden los vectores-columna. Designaremos mediante U (V) una matriz cuyas columnas son los vectores $u_1, \ldots, u_n$ $(v_1, \ldots, v_n)$. En tales circunstancias el hecho de que estos vectores satisfacen las correlaciones (107.2) significa que la matriz

$$C = V^* A U$$

es regular triangular izquierda. De aquí se desprende la siguiente descomposición de la matriz A en factores:

$$A = V^{-1*} C U^{-1}. \tag{107.8}$$

Así pues, el hecho de que se conocen los sistemas de vectores, A-seudoduales, salvo una normalización, permite resolver el sistema de ecuaciones algebraicas lineales (101.7) con la evaluación de errores (107.7) y, por otra parte, obtener la descomposición (107.8) de la matriz A en factores, entre los cuales hay uno triangular. Demostremos que la afirmación recíproca es también cierta. A saber, cualquier método de resolución de los sistemas de ecuaciones algebraicas lineales, basado en una descomposición de la matriz en factores, entre los cuales hay aunque sea un solo factor triangular, determina ciertos sistemas de vectores, A-seudoduales, salvo una normalización. Por consiguiente, al realizar estos métodos según los esquemas (107.3)—(107.6), se pueden utilizar las evaluaciones (107.7).

Examinemos la matriz P que se obtiene de una matriz unidad intercambiando sus columnas o, lo que es igual, intercambiando las filas en el orden inverso. Es fácil comprobar que la multiplicación a la derecha de la matriz C por P cambia de lugar en el orden inverso las columnas de la matriz C; la multiplicación a la izquierda de la matriz CP por P conmuta en el orden inverso las filas de la matriz CP. Por esta razón, los elementos f_{ij} de la matriz $F = PCP$ están vinculados con los elementos c_{ij} de la matriz C mediante una correlación $f_{ij} = c_{n-i+1\ n-j+1}$.

De aquí pueden sacarse toda una serie de corolarios útiles. Numeremos las diagonales de la matriz, paralelas a la principal, de abajo arriba, una tras otra, con los números $-(n-1), -(n-2), \ldots$ $\ldots, 0, \ldots, (n-2), (n-1)$. La diagonal con el número 0 será principal. Aceptada tal numeración, los elementos de la k-ésima diagonal se determinan por la correlación $j - i = k$. Si la matriz C satisface las condiciones $c_{ij} = 0, \quad k < j - i, \quad j - i < l$ para ciertos números $l \leqslant k$, entonces para la matriz $F = PCP$ se cumplirán las igualdades $f_{ij} = 0, \quad -l < j - i, \quad j - i < -k$. Por consiguiente, en la transformación de $F = PCP$ la matriz diagonal sigue siendo diagonal, la matriz triangular derecha (izquierda) se convierte en triangular izquierda (derecha) y la matriz bidiagonal derecha (izquierda), en bidiagonal izquierda (derecha), etc.

Supongamos ahora que se aplica cierto método de resolución del sistema de ecuaciones algebraicas lineales (107.1), basado en la descomposición previa de la matriz A en factores:

$$A = QCR, \tag{107.9}$$

donde la matriz C es triangular. Sin restringir esencialmente la generalidad, podemos considerar que C es triangular izquierda puesto que de lo contrario, en lugar de la descomposición (107.9) consideraríamos la descomposición

$$A = (QP)\,(PCP)\,(PR),$$

donde la matriz PCP debe ser triangular izquierda, de acuerdo con lo dicho más arriba. Las matrices buscadas U, V, que determinan los sistemas de vectores, A-seudoduales, salvo la normalización, $u_1, \ldots, u_n$ y $v_1, \ldots, v_n$, pueden definirse por las igualdades

$$U = R^{-1}, \qquad V = Q^{-1*}.$$

Hemos de notar que en la descomposición (107.9), engendrada por un método numérico, las matrices Q y R, son como regla, suficientemente sencillas. Estas son, con frecuencia, matrices unitarias o rectangulares y también las matrices que difieren de las triangulares en que las filas y columnas están permutadas. Por ello, las matrices R^{-1} y Q^{-1*} se hallan sin dificultad. En todo caso, los esfuerzos totales de cálculo para su determinación son considerablemente inferiores a aquellos que son necesarios para obtener la descomposición (107.9). Estas propiedades las poseen tales métodos ampliamente conocidos como el método de Gauss, de raíces cuadradas, de Jordan, de ortogonalización, de reflexiones, de giro, etc.; los métodos basados en la reducción del sistema a la forma bidiagonal con obtención de las descomposiciones normalizadas; los métodos de direcciones conjugadas, etc.

De este modo, la mayoría de los métodos numéricos que se emplean para la resolución de las ecuaciones operacionales (107.1) en un espacio de dimensión finita son, en realidad, métodos de construcción de los sistemas A-seudoduales de vectores. Pese a la diversidad de las formas concretas de los propios métodos, todos ellos pueden investigarse a partir de las posiciones generales, basándose siempre en el teorema 107.1.

Ejercicios.

1. Convengamos en considerar que un operador es una matriz y los vectores de un espacio, vectores columna. Demuéstrese que, aplicando el método de direcciones duales, los errores consecutivos están ligados entre sí por la correlación

$$e_k = (E - S_k)\, e_{k-1},$$

donde

$$S_k = \frac{u_k v_k^* A}{v_k^* A u_k}. \tag{107.10}$$

2. Demuestrese que los operadores S_k satisfacen las igualdades

$$S_k^2 = S_k, \qquad S_i S_k = 0,$$

$$S_i (E - S_k)(E - S_{k-1}) \ldots (E - S_1) = 0, \qquad i < k.$$

3. Demuéstrese que los operadores S_k de (107.10) y el operador P_k de (107.7) están ligados entre sí por la correlación

$$P_k = (E - S_k)(E - S_{k-1}) \ldots (E - S_1).$$

4. ¿Qué significan los operadores S_k y P_k para los métodos concretos, determinados por la descomposición (107.9)?

5. ¿Cómo varían los errores e_k para los métodos concretos, determinados por la descomposición (107.9)?

6. ¿Cuál de los métodos conocidos de resolución de los sistemas de ecuaciones algebraicas lineales no está basado en la descomposición (107.9)?

CONCLUSIÓN

En el presente manual al lector se le expone un material, suficientemente amplio, indispensable para comprender tanto la base teórica del álgebra lineal, como los métodos numéricos de éste. No obstante, debido a las peculiaridades de diferentes programas de estudio y a la escasez eventual del tiempo que se designa a las conferencias referentes a este curso, algunos apartados de este libro pueden escaparse de la atención del lector. Es por esta razón que damos a conocer aquí la característica general de todo el curso en consideración.

El álgebra lineal, como ciencia, estudia los conjuntos de estructura especial, así como las funciones que actúan en dichos conjuntos. En general, los problemas análogos se tratan también en otras ramas de las matemáticas, por ejemplo, en el análisis matemático. La característica singular del álgebra lineal radica precisamente en que los conjuntos son siempre espacios lineales de dimensión finita y las funciones intervienen como operadores lineales.

A la consideración de las propiedades generales de los espacios linales están dedicados los §§ 10, 13—21, las propiedades generales de los operadores lineales se examinan en los §§ 56—61, 63—74. La información, que se da en estos párrafos, puede obtenerse por varios métodos, incluso directamente sin recurrir a los conceptos y medios de investigación especiales, excepto los más elementales. Pero, uno de los conceptos adicionales merece una atención especial. Nos referimos al determinante.

Como función numérica, definida sobre los sistemas de vectores, el determinante es un ente relativamente sencillo. Sin embargo, posee varias propiedades de importancia. Estas propiedades lo convirtieron en un instrumento de amplio uso que facilita considerablemente la realización de diferentes investigaciones. Además, el determinante se emplea frecuentemente en la construcción de los métodos numéricos. Todo esto nos obligó a prestar mucha atención al concepto de determinante, al considerar en los §§ 34—42, 62 sus propiedades geométricas y algebraicas. Como instrumento de investigación el determinante se utiliza en el libro para demostrar las más diversas afirmaciones.

Otra función numérica de dos argumentos vectoriales —que es producto escalar— determina las dos clases más importantes de los espacios lineales, llamados euclídeos y unitarios. La nueva noción fundamental en dichos espacios es la de ortogonalidad. En los §§ 27—33 se estudian las propiedades de los espacios lineales, adicionales respecto del producto escalar, en los §§ 75—81, las propiedades de los operadores lineales, adicionales respecto del producto escalar.

Los sistemas de ecuaciones algebraicas lineales son de importancia exclusiva en toda la matemática, no sólo en el algebra lineal. Al estudio de diferentes aspectos de los sistemas mencionados están dedicados los §§ 22, 45, 46, 48.

Como regla, sólo el material citado constituye la base del curso del álgebra lineal al cual se añade, como un curso independiente, el de la geometría analítica. En el presente manual la información indispensable propia al curso de la geometría analítica se da no de modo aislado, sino en conjunto con la información correspondiente del álgebra lineal. La exposición semejante del material nos permitió lograr ciertas ventajas, a saber, se redujeron muchas demostraciones de un mismo tipo en ambos cursos, se consiguió subrayar la interpretación geométrica de tales conceptos algebraicos abstractos como el espacio lineal, el plano en un espacio lineal, el determinante, sistemas de ecuaciones algebraicas lineales, etc.

El álgebra lineal se enriquece en grado considerable con hechos nuevos, si en los espacios lineales se introducen las nociones sobre la distancia entre los vectores y límite de una sucesión de vectores. La necesidad de tal introducción está dictada también por las exigencias de los métodos numéricos. Las propiedades métricas de los espacios lineales se estudian en los §§ 49—54, las de los operadores lineales, en los §§ 82—84. Desde luego, todos estos problemas se dan comúnmente en el análisis funcional, pero, no se subrayan, como regla, muchos resultados, muy importantes para los espacios lineales de dimensión finita.

La resolución numérica de los problemas del álgebra lineal va acompañada casi siempre de la aparición de los errores de redondeo. Por esto el personal encargado de dos cálculos debe darse cuenta de los cambios en las propiedades de diferentes entes del álgebra lineal a que llevan las pequeñas variaciones de los vectores y operadores. A la influencia de las pequeñas perturbaciones están dedicados los §§ 33, 87, 89.

Las propiedades de muchos entes del álgebra lineal pueden cambiarse por otras contrarias incluso cuando las perturbaciones sean pequeñas. Así por ejemplo, un sistema linealmente dependiente de vectores puede convertirse en un sistema independiente o aumentar su rango, un operador de estructura de Jordan puede pasar a ser un operador de estructura simple y un sistema compatible de ecuaciones algebraicas lineales, un sistema incompatible, etc. Todos estos hechos originan dificultades muy grandes en la resolución práctica de los problemas.

Nuestra insistente sugerencia es que el lector lea una vez más con mucha atención el § 22 y analice el ejemplo que allí se aduce. El lector debe también reflexionar profundamente sobre las preguntas al final del párrafo.

A pesar de la inestabilidad de muchas nociones del álgebra lineal, sus problemas pueden resolverse de modo bien estable. Con el fin de demostrar tal posibilidad, se ha incluido en el libro la descripción de un método estable de resolución de los sistemas de ecuaciones algebraicas lineales. El fundamento teórico y el esquema general del método citado se dan en los §§ 85, 86, 88.

La parte final del libro viene dedicada a la descripción e investigación de diferentes problemas referentes a las formas bilineales y cuadráticas. Estas funciones numéricas son de mucha importancia en el álgebra lineal y están estrechamente ligadas con la construcción de los métodos numéricos. En los §§ 90—94 se consideran las propiedades generales de las formas bilineales y la relación de sus transformaciones con las descomposiciones matriciales; en los §§ 98—101, la ampliación de la noción de ortogonalidad; en los §§ 103—107 se muestra cómo se deben emplear las formas bilineales en los procesos de cálculo.

INDICE ALFABETICO