1er Coloquio del Departamento de Matemáticas

El Teorema Espectral en Dimensión Finita

Gustavo Nicolás Izquierdo Buenrostro

Casa abierta al tiempo

El Teorema Espectral en Dimensión Finita

Gustavo Nicolás Izquierdo Buenrostro

Departamento de Matemáticas, UAM-I

Casa abierta al tiempo

Universidad Autónoma Metropolitana

Contents

Capítulo 1

Producto interior y espacios de dimensión finita

1.1. Trabajo y producto punto

El producto interior aparece de manera natural en problemas de fisica asociados con el concepto de trabajo. Así pues, empezaremos por dar una breve discusión sobre como se calcula el trabajo. Si se recuerda, en una primera instancia el concepto físico de trabajo se define como el producto de la magnitud de la fuerza por la distancia recorrida al aplicar dicha fuerza. Sin embargo, cuando se conciben las fuerzas y los desplazamientos como vectores en el plano o en el espacio, la definición de trabajo debe tomar en cuenta que las fuerzsa y el deplazamiento resultante no necesariamente son paralelos.

En este caso la definición del trabajo es el producto de la componente de la fuerza en la dirección del movimiento por la distancia recorrida.

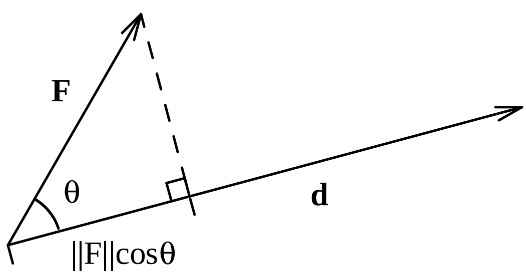

Figure 1.1

De la trigonometría elemental uno tiene que la componente de la fuerza en la dirección del desplazamiento es $\|\mathbf{F}\| \cos\theta$, donde es el ángulo θ que forman la fureza $\mathbf{F}$ y el desplazamiento $\mathbf{d}$ (ver figura 1.1) y $\|\mathbf{F}\|$ es la magnitud de la fuerza. Como la distancia recorrida es la magnitud del desplazamiento, esto es $\|\mathbf{d}\|$ se obtiene que el trabajo físico T está dado por la fórmula

$$T = \left(\|\mathbf{F}\| \cos\theta\right) \|\mathbf{d}\| = \|\mathbf{F}\| \, \|\mathbf{d}\| \cos\theta$$

Así, el trabajo se puede calcular si se conoce el ángulo θ que forman $\mathbf{F}$ y $\mathbf{d}$ y sus correspondientes magnitudes.

A partir de esta fórmula se puede obtener una expresión para el trabajo en términos de las componentes de $\mathbf{F}$ y $\mathbf{d}$ en un sistema rectangular en el plano o el espacio. En efecto, si las componentes de la fuerza y el desplazamiento están dadas por

$$\begin{aligned} \mathbf{F} &= f_1\mathbf{i} + f_2\mathbf{j} + f_3\mathbf{k} = (f_1, f_2, f_3) \\ \mathbf{d} &= d_1\mathbf{i} + d_2\mathbf{j} + d_3\mathbf{k} = (d_1, d_2, d_3) \end{aligned}$$

Entonces, las magnitudes de $\mathbf{F}$ y $\mathbf{d}$ son

$$\|\mathbf{F}\| = \sqrt{f_1^2 + f_2^2 + f_3^2} \quad \text{y} \quad \|\mathbf{d}\| = \sqrt{d_1^2 + d_2^2 + d_3^2}$$

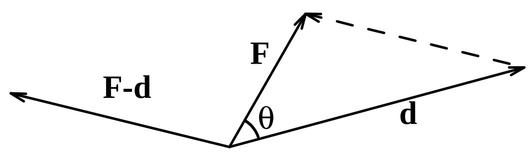

Figure 1.2

Ahora bien, para obtener una expresión que involucre la expresión $\|\mathbf{F}\|\|\mathbf{d}\| \cos\theta$ y las componentes rectangulares de $\mathbf{F}$ y $\mathbf{d}$ se usa la *ley de los Cosenos* en el triángulo que forman $\mathbf{F}$, $\mathbf{d}$ y el segmento que une el extremo final de $\mathbf{d}$ con el de $\mathbf{F}$ (ver Figura 2). Puesto que el segmento que une a $\mathbf{d}$ con $\mathbf{F}$ es paralelo y de la misma magnitud que el vector $\mathbf{F} - \mathbf{d}$ se sigue de la ley de los cosenos que

$$\|\mathbf{F} - \mathbf{d}\|^2 = \|\mathbf{F}\|^2 + \|\mathbf{d}\|^2 - 2\|\mathbf{F}\| \, \|\mathbf{d}\| \cos\theta$$

Si ahora expresamos las normas de **F**, **d** y **F** − **d** en terminos de sus componentes nos queda la relación

$$(f_1 - d_1)^2 + (f_2 - d_2)^2 + (f_3 - d_3)^2 \\ = \left(f_1^2 + f_2^2 + f_3^2\right) + \left(d_1^2 + d_2^2 + d_3^2\right) - 2\,\|\mathbf{F}\|\,\|\mathbf{d}\|\cos\theta$$

Al desarrollar la expresión del lado izquierdo y cancelar términos nos queda que

$$f_1 d_1 + f_2 d_2 + f_3 d_3 = \|\mathbf{F}\|\,\|\mathbf{d}\|\cos\theta \tag{1.1}$$

y por lo tanto el trabajo realizado por una fuerza **F** con componentes (f_1, f_2, f_3) que producen un despalzamiento $\mathbf{d} = (d_1, d_2, d_3)$ está dado por

$$T = f_1 d_1 + f_2 d_2 + f_3 d_3$$

Como el trabajo se define originalmente como un producto, la expresión del lado derecho de esta igualdad se suele representar como $\mathbf{F} \cdot \mathbf{d}$, esto es

$$\mathbf{F} \cdot \mathbf{d} = f_1 d_1 + f_2 d_2 + f_3 d_3$$

y es llamado el *producto punto* de **F** por **d**.

1.2. El producto punto y la geometría euclidiana en el espacio

La simplicidad para calcular $\mathbf{F} \cdot \mathbf{d}$ en términos de componentes rectangulares junto con la igualdad

$$\mathbf{F} \cdot \mathbf{d} = f_1 d_1 + f_2 d_2 + f_3 d_3 = \|\mathbf{F}\|\,\|\mathbf{d}\|\cos\theta \tag{1.2}$$

le da al producto punto un carácter geométrico relevante ya que los dos elementos fundamentales de la geometría euclidiana, magnitud y ángulo, se sintetizan en esta expresión.

Puesto que las fuerzas y desplazamientos son vectores y sus magnitudes corresponden a la noción de longitud euclidiana de vectores, podemos abstraer la definición de producto punto de su contenido físico y estudiar sus propiedades geométricas y algebráicas. Así, de manera más general, nosotros damos la siguiente definición:

Definición 1.2.1. Dados $\vec{x}$ y $\vec{y}$, vectores en $\mathbb{R}^3$ con componentes (x_1, x_2, x_3) y $(y_1, y_2 y_3)$, definimos el producto punto de $\vec{x}$ por $\vec{y}$, $\vec{x} \cdot \vec{y}$ como

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

En estos términos la relación (1.2) toma la forma

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 = \|\vec{x}\|\,\|\vec{y}\|\cos\theta \tag{1.3}$$

donde θ es el ángulo que forman $\vec{x}$ y $\vec{y}$.

Una primera consecuencia de esta igualdad es que si $\vec{x}$ y $\vec{y}$ son el mismo, entonces el ángulo que forman es 0 y como $\cos 0 = 1$ se tiene que

$$\vec{x} \cdot \vec{x} = \|\vec{x}\| \, \|\vec{x}\| = \|\vec{x}\|^2 \tag{1.4}$$

y por ende

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}}$$

Otra manera de llegar a esto es observando que de acuerdo a la definición $\vec{x} \cdot \vec{x} = x_1^2 + x_2^2 + x_3^2$ y, en consecuencia,

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{\vec{x} \cdot \vec{x}}$$

También es posible calcular el ángulo entre dos vectores diferentes de cero usando el producto punto. De la ecuación (1.3) se sigue que

$$\cos \theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}$$

y por lo tanto

$$\theta = \arccos \left(\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|} \right) \tag{1.5}$$

Puesto que los dos elementos fundamentales de la geometría euclidiana, longitud y ángulo, se sintetizan en al producto interior, varios conceptos geométricos pueden formularse en términos del producto punto. Así, por ejemplo, del hecho de que dos vectores son paralelos si el ángulo que forman es de 0° o 180° grados uno obtiene de (1.5) con $\theta = 0^\circ$ o 180° que $\vec{x} \cdot \vec{y} = (\pm 1) \|\vec{x}\| \, \|\vec{y}\|$ (o en forma equivalente $|\vec{x} \cdot \vec{y}| = \|\vec{x}\| \, \|\vec{y}\|$). Más aún, el recíproco tambien se cumple: De la igualdad (1.5) se tiene que si $\vec{x}$ y $\vec{y}$ son diferentes de cero y $\vec{x} \cdot \vec{y} = (\pm 1) \|\vec{x}\| \, \|\vec{y}\|$, entonces,

$$\theta = \arccos \left(\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|} \right) = \arccos (\pm 1)$$

y por lo tanto el ángulo entre $\vec{x}$ y $\vec{y}$ es 0° o 180°, esto es, los vectores son paralelos. En síntesis

Dos vectores en $\mathbb{R}^3$ diferentes de cero $\vec{x}$ y $\vec{y}$ son paralelos, si y solo si,

$$|\vec{x} \cdot \vec{y}| = \|\vec{x}\| \, \|\vec{y}\|$$

De modo similar el concepto de perpendicularidad también se puede caracterizar en términos del producto punto. En este caso sabemos que cuando dos vectore son perpendiculares el ángulo que forman es 90° y como $\cos 90^\circ = 0$ se tiene de (1.3) que $\vec{x} \cdot \vec{y} = 0$ y recíprocamente para vectores diferentes de cero se sigue de (1.5) que si $\vec{x} \cdot \vec{y} = 0$, enmtonces, el ángulo entre $\vec{x}$ y $\vec{y}$ es 90° y en consecuencia

Dos vectores en $\mathbb{R}^3$ diferentes de cero $\vec{x}$ y $\vec{y}$ son perpendiculares, si y solo si

$$\vec{x} \cdot \vec{y} = 0$$

También podemos reformular el concepto de distancia. Si x y y son dos puntos con vectores de posición $\vec{x}$ y $\vec{y}$ respectivamente, entonces, el segmento que una a x con y es de la misma longitud y dirección que el vector $\vec{y} - \vec{x}$ (ver Figura 3) pero esta longitud es justamente $\|\vec{x} - \vec{y}\|$

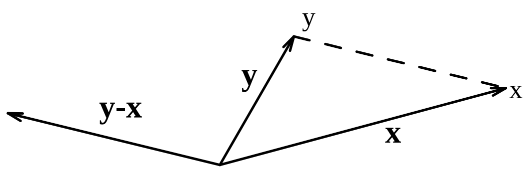

Figure 1.3

Por lo tanto la distancia entre dos elementos de $\mathbb{R}^3$, $\vec{x}$ y $\vec{y}$ está dada por

$$\text{dist}\left(\vec{x}, \vec{y}\right) = \|\vec{x} - \vec{y}\|$$

Admás de esto, el el producto punto tiene las siguientes propiedades:

PROPOSICIÓN 1.2.2. *Para cualesquiera vectores $\vec{x}, \vec{y}$ y $\vec{z}$ en $\mathbb{R}^3$ se tiene*

i) $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ *y* $\left(\vec{x} + \vec{y}\right) \cdot \vec{z} = \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}$

ii) $\vec{x} \cdot (\lambda \vec{y}) = \lambda(\vec{x} \cdot \vec{y}) = \left(\lambda \vec{x}\right) \cdot \vec{y}, \lambda \in \mathbb{R}$

iii) $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$

iv) $\vec{x} \cdot \vec{x} \geq 0$ *y* $\vec{x} \cdot \vec{x} = 0 \Leftrightarrow \vec{x} = \vec{0}$.

v) $\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2$ *y* $\|\vec{x} - \vec{y}\|^2 = \|\vec{x}\|^2 - 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2$

DEMOSTRACIÓN. Mostraremos solo la primera parte de las propiedades i) y v) el resto se deja como ejercicio.

i) Si $\vec{x}, \vec{y}$ y $\vec{z}$ son vectores en $\mathbb{R}^3$ con componentes (x_1, x_2, x_3), (y_1, y_2, y_3) y (z_1, z_2, z_3), respectivamente, entonces

$$\vec{y} + \vec{z} = (y_1 + z_1, y_2 + z_2, y_3 + z_3)$$

y de la definición se tiene que

$$\begin{aligned}\vec{x}\cdot\left(\vec{y}+\vec{z}\right) &= x_1\left(y_1+z_1\right)+x_2\left(y_2+z_2\right)+x_3\left(y_3+z_3\right)\\ &= \left(x_1y_1+x_2y_2+x_3y_3\right)+\left(x_1z_1+x_2z_2+x_3z_3\right)\end{aligned}$$

pero $x_1y_1+x_2y_2+x_3y_3=\vec{x}\cdot\vec{y}$ y $x_1z_1+x_2z_2+x_3z_3=\vec{x}\cdot\vec{z}$ por lo tanto

$$\vec{x}\cdot\left(\vec{y}+\vec{z}\right)=\vec{x}\cdot\vec{z}$$

v) Para la demostración asumimos que las propiedades i) a iv) ya se han demostrado. Ahora bien, de (1.4) se tiene que

$$\|\vec{x}+\vec{y}\|^2=\left(\vec{x}+\vec{y}\right)\cdot\left(\vec{x}+\vec{y}\right)$$

De la propiedad i) y de (1.4) obtenemos que

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &= \left(\vec{x}+\vec{y}\right)\cdot\left(\vec{x}+\vec{y}\right)\\ &= \left(\vec{x}+\vec{y}\right)\cdot\vec{x}+\left(\vec{x}+\vec{y}\right)\cdot\vec{y}\\ &= \vec{x}\cdot\vec{x}+\vec{y}\cdot\vec{x}+\vec{x}\cdot\vec{y}+\vec{y}\cdot\vec{y}\\ &= \|\vec{x}\|^2+\vec{y}\cdot\vec{x}+\vec{x}\cdot\vec{y}+\|\vec{y}\|^2\end{aligned}$$

Fimalmente de iii) sabemos que $\vec{y}\cdot\vec{x}=\vec{x}\cdot\vec{y}$ y por ende

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &= \|\vec{x}\|^2+\vec{y}\cdot\vec{x}+\vec{x}\cdot\vec{y}+\|\vec{y}\|^2\\ &= \|\vec{x}\|^2+2\vec{x}\cdot\vec{y}+\|\vec{y}\|^2\end{aligned}$$ □

Las propiedades i) a iii) se pueden sintetizar diciendo que le producto punto visto como función de $\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}$ es una forma bilineal simétrica.En particular uno tiene que

$$x\cdot(\beta\vec{y}+\gamma\vec{z})=\beta\vec{x}\cdot\vec{y}+\gamma\vec{x}\cdot\vec{y} \text{ y } \left(\alpha\vec{x}+\beta\vec{y}\right)\cdot\vec{z}=\alpha\vec{x}\cdot\vec{z}+\beta\vec{y}\cdot\vec{z}$$

Otras propiedades de caracter geométrico del producto punto son las siguientes

Proposición 1.2.3. *Sean $\vec{x}$ y $\vec{y}$ vectores en $\mathbb{R}^3$, entonces,*

i) (Teorema de Pitágoras) Si $\vec{x}$ y $\vec{y}$ son perpendiculares, entonces

$$\|\vec{x}+\vec{y}\|^2=\|\vec{x}\|^2+\|\vec{y}\|^2 \ \ y \ \ \|\vec{x}-\vec{y}\|^2=\|\vec{x}\|^2-\|\vec{y}\|^2$$

ii) (El único vector perpendicular a cualquier otro es el $\vec{0}$) Si $\vec{x}\cdot\vec{y}=0$ para todo $\vec{y}\in\mathbb{R}^3$, necesariamente $\vec{x}=\vec{0}$.
iii) (Ley de los cosenos) $\|\vec{x}-\vec{y}\|^2=\|\vec{x}\|^2+\|\vec{y}\|^2-2\vec{x}\cdot\vec{y}$
iv) (Identidad de Paralelogramo) $\|\vec{x}+\vec{y}\|^2+\|\vec{x}-\vec{y}\|^2=2(\|\vec{x}\|^2+\|\vec{y}\|^2)$.

Estos resultados y muchas otras propiedades geométrica se siguen de la relació (1.3) y de la proposición anterior (ver los ejercicios 1 a 8 de este capítulo).

1.3. El producto punto en $\mathbb{R}^N$

La relevancia geométrica del producto punto estriba en el hecho de que esta operación junto con las propiedades de espacio vectorial nos permiten recuperar todas las propiedades de la geometría euclideana. Nuestra definición de producto punto es susceptible de ser generalizada a $\mathbb{R}^n$. Nosotros intentaremos llevar adelante esta generalización, procurando destacar las propiedades básicas que le dan ese caracter geométrico al produnto punto.Consideremos el espacio de todas las "eneadas" de números reales, con la suma y multiplicación por escalares definidos de la manera usual. Esto es, el espacio vectorial formado por los vectores de la forma $\vec{x} = (x_1, \ldots, x_n)$, $x_i \in \mathbb{R}$, $i = 1, \ldots, n$, con la suma

$$\vec{x} + \vec{y} = (x_1 + y_1, ..., x_n + y_n)$$

y la multiplicación por escalares

$$\lambda\vec{x} = (\lambda x_1, \ldots, \lambda x_n)$$

Este espacio es llamado $\mathbb{R}^n$.

La generlización natural del producto punto a vectores en $\mathbb{R}^n$ es la siguiente

Definición 1.3.1. Definimos el producto punto de dos vectores en $\mathbb{R}^n$, $\vec{x} = (x_1, \ldots, x_n)$ y $\vec{y} = (y_1, \ldots, y_n)$, como

$$\vec{x} \cdot \vec{y} = x_1 y_1 + ... + x_n y_n$$

El especio vectorial $\mathbb{R}^n$ junto con el producto punto es llamado el **espacio euclidiano de dimensión n** y se denota por E^n

A partir de esta definición y las propiedades algebraicas de los vectores en $\mathbb{R}^n$ uno puede mostrar la siguiente versión de la Proposición 1.2.2:

Proposición 1.3.2. *Para cualesquiera vectores $\vec{x}$, $\vec{y}$ y $\vec{z}$ en $\mathbb{R}^n$ se tiene*

i) $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ *y* $\left(\vec{x} + \vec{y}\right) \cdot \vec{z} = \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}$

ii) $\vec{x} \cdot (\lambda\vec{y}) = \lambda(\vec{x} \cdot \vec{y}) = \left(\lambda\vec{x}\right) \cdot \vec{y}, \lambda \in \mathbb{R}$

iii) $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$

iv) $\vec{x} \cdot \vec{x} \geq 0$ *y* $\vec{x} \cdot \vec{x} = 0 \Leftrightarrow \vec{x} = \vec{0}$.

La demostración de estas propiedadeses exactamente igual que la de la Proposición 1.2.2 excepto que en este caso se consideran n componentes en lugar de 3.

De acuerdo con la relación (1.4) tendemos que

$$\vec{x} \cdot \vec{x} = x_1^2 + x_2^2 + \cdots + x_n^2$$

por la tanto la manera de clcular la longitud de un vector en el espacio euclidiano E^n tendrá que ser la siguiente:

DEFINICIÓN 1.3.3. Dado un vector $\vec{x} = (x_1, \ldots, x_n)$ en E^n la **norma, magnitud o longitud euclidiana** del vector $\vec{x}$ está dada por

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$$

y definimos l**a distancia entre dos vectores** $\vec{x}$ y $\vec{y}$ de E^n como

$$\text{dist}\left(\vec{x}, \vec{y}\right) = \|\vec{y} - \vec{x}\|$$

Notemos que la propiedad iv) de la Proposición 1.3.2 garantiza que la norma euclidiana está definida para todo vector en E^n

A partir de esta noción de longitud o norma y la Proposición 1.3.2 uno obtiene las siguientes propiedades:

PROPOSICIÓN 1.3.4. *Sean $\vec{x}$ y $\vec{y}$ vetores en E^n,entonces,*

i) $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$
ii) $\|\vec{x}\| \geq 0$ *y* $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$
iii) $\|\lambda\vec{x}\| = |\lambda| \, \|\vec{x}\|$
iv) $\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2$ *y* $\|\vec{x} - \vec{y}\|^2 = \|\vec{x}\|^2 - 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2$
v) $4\vec{x} \cdot \vec{y} = \|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2$

DEMOSTRACIÓN. i) Se sigue directamente de la definición.

ii) Puesto que en la definición de norma se considera la raíz positiva de $\vec{x} \cdot \vec{x}$, es claro que $\|\vec{x}\| \geq 0$. Por otra parte sabemos que $\vec{x} \cdot \vec{x} = 0$ si y solo si $\vec{x} = 0$ (prop 1.3.2-iv), por lo tanto $\|\vec{x}\|^2 = 0 \Leftrightarrow \vec{x} = \vec{0}$ que es equivalente a la segunda parte de iv)

iii) De la definición de norma y la proposición 1.3.2-ii) se sigue que

$$\begin{aligned} \|\lambda\vec{x}\| &= \sqrt{\left(\lambda\vec{x}\right) \cdot \left(\lambda\vec{x}\right)} = \sqrt{\lambda\left(\vec{x} \cdot \left(\lambda\vec{x}\right)\right)} \\ &= \sqrt{\lambda^2\left(\vec{x} \cdot \vec{x}\right)} = |\lambda| \sqrt{\left(\vec{x} \cdot \vec{x}\right)} \\ &= |\lambda| \, \|\vec{x}\| \end{aligned}$$

lo que muestera iii).

iv) Mostraremos solo la segunda parte. De las propiedades de espacio vectorial de E^n se tiene que

$$\|\vec{x} - \vec{y}\|^2 = \left\|\vec{x} + \left((-1)\,\vec{y}\right)\right\|^2$$

y de i) obtenemos

$$\left\|\vec{x} + \left((-1)\,\vec{y}\right)\right\|^2 = \left(\vec{x} + \left((-1)\,\vec{y}\right)\right) \cdot \left(\vec{x} + \left((-1)\,\vec{y}\right)\right)$$

finalmente combinando i) y ii) de la Proposición 1.3.2

$$\begin{aligned}\|\vec{x}-\vec{y}\|^2 &= \left(\vec{x}+\left((-1)\vec{y}\right)\right)\cdot\left(\vec{x}+\left((-1)\vec{y}\right)\right)\\ &= \vec{x}\cdot\vec{x}+(-2)\vec{x}\cdot\vec{y}+\vec{y}\cdot\vec{y}\\ &= \|\vec{x}\|^2-2\vec{x}\cdot\vec{y}+\|\vec{y}\|^2\end{aligned}$$ □

En la sección anterior vimos que el producto punto en el espacio está vinculado con la nocón de ángulo, mediante la fórmula $\vec{x}\cdot\vec{y} = \|\vec{x}\|\,\|\vec{y}\|\cos\theta$. Sin embargo, en $\mathbb{R}^n$ no hay una manera natural de definir el ángulo entre dos vectores. Una alternativa en este caso es definir el ángulo entre dos vectores diferentes de cero $\vec{x}$ y $\vec{y}$ vía la fórmula

$$\theta = \arccos\left(\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}\right)$$

La dificultad de esta definición estriba en que el arcocoseno esta definido para valores entre -1 y 1 y por ende para que el ángulo esté bien definido para cualesquiera dos vectores diferentes de cero en $\mathbb{R}^n$ es necesario que el cociente $\vec{x}\cdot\vec{y}/\,\|\vec{x}\|\,\|\vec{y}\|$ sea un valor entre -1 y 1. Afortunadamente esta condición es garantizada por la famosa desigualdad de Cauchy- Buniakovski- Schwarz.

Proposición 1.3.5 (Cauchy-Buniakovski-Schwarz). *Sean $\vec{x}$ y $\vec{y}$ vectores en E^n, entonces,*

$$|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\,\|\vec{y}\|$$

Más aún, la igualdad se da si y sólo si $\vec{x} = \lambda\vec{y}$ para algún número real λ.

Demostración. Es claro que si $\vec{y} = 0$ la desigualda se cumple, así pues asumiremos que $\vec{y} \neq 0$. De la propiedad iv) uno tiene que

$$0 \leq \|\vec{x}-\mu\vec{y}\|^2 = (\vec{x}-\mu\vec{y})\cdot(\vec{x}-\mu\vec{y}), \forall\mu\in\mathbb{R}$$

Utilizando las propiedades i) y iii) obtenemos

$$0 \leq \|\vec{x}\|^2 - 2\mu\vec{x}\cdot\vec{y} + \mu^2\|\vec{y}\|^2, \forall\mu\in\mathbb{R}$$

Puesto que esta desigualdad es válida para toda μ real, también es válida para el valor de μ para el cual la expresión $\|\vec{x}\|^2 - 2\mu\vec{x}\cdot\vec{y} + \mu^2\,\|\vec{y}\|^2$ toma su valor mínimo, esto ocurre cuando $\mu = \vec{x}\cdot\vec{y}/\,\|\vec{y}\|^2$. De esto se sugue la desigualdad $0 \leq \|\vec{x}\|^2 - (\vec{x}\cdot\vec{y})^2/\,\|\vec{y}\|^2$. Despejando el producto punto obtenemos la desigualdad deseada.

Para ver que la igualdad se cumple si y sólo si los vectores son paralelos, empecemos por obsevar que en el caso en el que $|\vec{x}\cdot\vec{y}| = \|\vec{x}\cdot\vec{y}\|$ obtenemos, para $\mu = \pm\|\vec{x}\|\,/\,\|\vec{y}\|$ que $\|\vec{x}-\mu\vec{y}\|^2 = 0$ lo que implica que $\vec{x} = \mu\vec{y}$. Recíprocamente, si $\vec{x} = \lambda\vec{y}$ se tiene del hecho de que $\|\lambda\vec{y}\|^2 = \lambda^2\,\|\vec{y}\|^2$ y de $\|\vec{x}-\lambda\vec{y}\|^2 = 0$ que $\vec{x}\cdot\vec{y} = \|\vec{x}\|\,\|\vec{y}\|$. □

OBSERVACIÓN 1. Es importante señalar que la demostración de la desigualdad se basa exclusivamente en las propiedades de i) a iv) de la Proposición 1.3.2 del producto punto.

Con base en este resultado nosotros damos la siguiente definición:

DEFINICIÓN 1.3.6. *El ángulo entre dos vectores no nulos $\vec{x}$ y $\vec{y}$ en E^n* se define como

$$\arccos\left(\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}\right)=\theta$$

donde arccos se considera como la rama de la función inversa del coseno con valores en $[0,\pi]$.

Tomando esta definición de ángulo como punto de partida y la norma euclidiana, nuevamente tenemos vía el producto punto la geometría euclidiana en E^n. En particular, cuando el producto punto de dos vectores es cero, se tiene que $\theta = \arccos(0) = \frac{\pi}{2}$ y por ende los vectores forman un ángulo recto. Esto da origen a la siguiente definición.

DEFINICIÓN 1.3.7. Dos vectores $\vec{x}$ y $\vec{y}$ son *perpendiculares* u *ortogonales* si y solo si

$$\vec{x}\cdot\vec{y}=0$$

En este caso escribiremos $\vec{x}\perp\vec{y}$.

Algunas de estas propiedades geométricas se dan en la siguiente

PROPOSICIÓN 1.3.8. *Sean $\vec{x}$ y $\vec{y}$ vectores en E^n, entonces,*

i) $\vec{x}\cdot\vec{y}=0$ *si y sólo si el ángulo entre $\vec{x}$ y $\vec{y}$ es $\frac{\pi}{2}$ (esto es, si y sólo si los vectores son perpendiculares)*
ii) *(Teorema de Pitágoras) Si el ángulo entre $\vec{x}$ y $\vec{y}$ es $\frac{\pi}{2}$, entonces,* $\|\vec{x}+\vec{y}\|^2=\|\vec{x}\|^2+\|\vec{y}\|^2$
iii) *Si $\vec{x}\cdot\vec{y}=0$ para todo $\vec{y}\in E^n$, necesariamente $\vec{x}=0$.*
iv) *(Ley de los cosenos)* $\|\vec{x}-\vec{y}\|^2=\|\vec{x}\|^2+\|\vec{y}\|^2-2\vec{x}\cdot\vec{y}$
v) *(Identidad de Paralelogramo)* $\|\vec{x}+\vec{y}\|^2+\|\vec{x}-\vec{y}\|^2=2(\|\vec{x}\|^2+\|\vec{y}\|^2)$.

Estas propiedades geométricas se dejan como ejercicios.

1.4. Espacios vectoriales y el Producto interior.

Como se observó en la sección anterior, las propiedades básicas del producto punto que permiten establecer la geometría euclidiana son las propiedades de la i) a iv) de la proposición 1. Mostraremos en esta sección que estas propiedades son la base para poder introducir una geometría euclidiana en un espacio vectorial real arbitrario.

DEFINICIÓN 1.4.1. Sea V un espacio vectorial real, una función $p\colon V \times V \to \mathbb{R}$ es llamada un *producto interior para* V si p satisface las siguientes propiedades:

i) $p(u, v + w) = p(u, v) + p(u, w)$
ii) $p(u, \lambda v) = \lambda p(u, v)$
iii) $p(u, v) = p(v, u)$
iv) $p(u, u) \geq 0, \forall u \in V$ y $p(u, u) = 0 \Leftrightarrow u = 0$

NOTACIÓN 1. Cuando una norma p se ha elegido como el producto interior en V, nosotros escribiremos $\langle u, v\rangle$ en lugar de $p(u, v)$ para remarcar que la forma p se considera como un producto interior en V.

Para introducir la geometría euclidiana en un espacio vectorial con un producto interior $\big(V, \langle\cdot,\cdot\rangle\big)$ es requisito indispensable que en V se considere como noción de longitud aquella definida por el producto interior. En términos más precisos

DEFINICIÓN 1.4.2. Sea $\big(V, \langle\cdot,\cdot\rangle\big)$ un espacio con producto interior. La *longitud euclidiana o norma dos* de un vertor u en V, con respecto al producto interior $\langle\cdot,\cdot\rangle$, se define como

$$\|u\|_2 = \sqrt{\langle u, u\rangle}$$

y definimos *la distancia entre dos elementos* u y v de V como

$$\operatorname{dist}(u, v) = \|v - u\|_2$$

Vemos ahora lgunos ejemplos de espacios vectoriales con producto interior

EJEMPLO 1.1. En $\mathbb{R}^2$ la forma $p(\vec{x}, \vec{y}) = 4x_1y_1 + 2(x_1y_2 + x_2y_1) + 4x_2$ define un producto interior. La longitud euclidiana con respecto aeste producto es $\|\vec{x}\|_2 = 2\sqrt{x_1^2 + x_1x_2 + x_2^2}$.

EJEMPLO 1.2. Denotemos por $P_n[x]$ el conjunto de polinomios de grado n, esto es, el conjunto de funciones de la forma $u(x) = a_0 + a_1x + \cdots + a_nx^n$ con $a_0, \ldots, a_n$ números reales. Este es un espacio vectorial con la suma y el producto por escalares de funciones. Dados los polinomios $u(x) = a_0 + a_1x + \cdots + a_nx^n$ y $v(x) = b_0 + b_1x + \cdots + b_nx^n$, definimos la forma $p_1(u, v) = \langle u, v\rangle_1$ como

$$\langle u, v\rangle_1 = p_1(u, v) = a_1b_1 + \cdots + a_nb_n = \sum_{i=0}^{n} a_ib_i$$

esta forma define un producto interior en $P_n[x]$. En este caso la norma dos es

$$\|u\|_{2,1} = \sqrt{\sum_{i=0}^{n} a_i^2}$$

EJEMPLO 1.3. Definimos $P_n[\cos x]$ como el conjunto de polinomios trigonométricos de al forma

$$u(x) = a_0 + a_1 \cos x + ... + a_n \cos^n x$$

y definimos $p_2 : P_n[\cos x] \times P_n[\cos x] \to \mathbb{R}$ como

$$\langle u, v \rangle_2 = p_2(u, v) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} u(x)v(x)dx$$

La forma p_2 define un producto interior en $P_n[\cos x]$ y la norma euclidiana correspondiente es

$$\|u\|_{2,2} = \sqrt{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (u\,(x))^2\, dx}$$

EJEMPLO 1.4. En $P_n[x]$ la forma $p_2(u, v) = \langle u, v \rangle_2$ dada por

$$\langle u, v \rangle_2 = \int_{-1}^{1} u(x)v(x)dx$$

también define un producto interior en $P_n[x]$ y

$$\|u\|_{2,2} = \sqrt{\int_{-1}^{1} (u\,(x))^2\, dx}$$

Notemos que en un mismo espacio vectorial se pueden definir diferentes productos interiores, cada uno de ellos, por supuesto, nos dará una manera de medir longitudes y ángulos, además de una interpretación de la geometría euclidiana en ese espacio. La decisión de cual producto interior deberá elegirse en un determinado espacio vectorial dependerá del tipo de problemas que se quieran resolver. Uno deberá tener siempre en mente que si se introduce un producto interior, entonces la manera de medir longitudes, esto es, la norma queda subordinada a dicho producto.

Como consecuencia de las definiciones de producto interior y norma dos se tienen los resultado ánalogos a las proposiciones 1.3.4 y 1.3.5.

TEOREMA 1.4.3. *La logitud euclidiana o norma en V dada por el producto interior $\langle \cdot,\cdot \rangle$ tiene las siguientes propiedades,*

i) $\|u\|_2 \geq 0$ *y* $\|u\|_2 = 0$ *si y sólo si* $u = 0$
ii) $\|\lambda u\|_2 = |\lambda|\, \|u\|_2$
iii) $\|u + v\|_2^2 = \|u\|_2^2 + 2\langle u, v \rangle + \|v\|_2^2$ *y* $\|u - v\|_2^2 = \|u\|_2^2 - 2\,\langle u, v \rangle + \|v\|_2^2$
iv) $4\,\langle u, v \rangle = \|u + v\|_2^2 - \|u - v\|_2^2$

TEOREMA 1.4.4 (Desigualdad de Cauchy-Bunikovski-Schwarz). *Sea $\langle\cdot,\cdot\rangle$ un producto interior en V, entonces,*

$$|\langle u, u\rangle| \leq \|u\|_2 \|v\|_2$$

Más aún la igualdad se da si y solo si $u = \lambda v$ para alguna λ real.

La demostración de estas propiedades es enteramente análoga a la de las proposiciones 1.3.4 y 1.3.5.

Puesto que en este caso también se cuenta con la desigualdad de Cauchy-Bunikovski-Schwarz podemos definir el concepto de ángulo en un espacio vectorial real con producto interior

DEFINICIÓN 1.4.5. Dado un espacio vectorial con producto interior $\left(V, \langle\cdot,\cdot\rangle\right)$ definimos el ángulo entre dos elementos no nulos u y v de V como

$$\theta = \arccos\left(\frac{u \cdot v}{\|u\|_2 \|v\|_2}\right)$$

Donde arccos toma valores en $[0, \pi]$.

Una consecuencia inmediata de esto es el siguiente resultado:

TEOREMA 1.4.6. *Sea V un espacio vectorial con el producto interior $\langle\cdot,\cdot\rangle$, entonces, para cualesquiera vectores diferentes de cero $u, v \in V$*

$$\langle u, v\rangle = \|u\|_2 \|v\|_2 \cos\theta$$

De modo análogo a como se hizo en la sección precedente tambien tenemos una manera de definir perpendicularidad entre vectores.

DEFINICIÓN 1.4.7. Sea $\left(V, \langle\cdot,\cdot\rangle\right)$ un espacio vectorial real con producto interior. Diremos que dos vectores u y v en V son **ortogonales** si y sólo si

$$\langle u, v\rangle = 0$$

En tal caso escribiremos $u \perp v$.

Por supuesto un puede demostrar la siguiente versión de la Proposición 1.3.8

PROPOSICIÓN 1.4.8. *Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio con producto interior y sean u y v vectores en V, entonces,*

i) $u \cdot v = 0$ *si y sólo si el ángulo entre u y v es $\frac{\pi}{2}$ (esto es, si y sólo si los vectores son perpendiculares)*

ii) *(Teorema de Pitágoras) Si el ángulo entre u y v es $\frac{\pi}{2}$, entonces,* $\|u+v\|^2 = \|u\|^2 + \|v\|^2$

iii) *Si $u \cdot v = 0$ para todo $v \in V$, necesariamente $u = 0$.*

iv) *(Ley de los cosenos)* $\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2u \cdot v$

v) *(Identidad de Paralelogramo)* $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$.

1.5. Proyección ortogonal

Un problrma de interés es el siguiente: Dado un subespacio y un punto v encontrar la distancia más corta del punto v al subespacio. Veremos en esta sección como vía el producto interior se puede encontrar el punto p del subespacio más cercano al punto v y, en consecuencia la distancia más corta de v al subespacio.

Consideremos primero el caso de una recta en el espacio, la cual es generada por un vector $\vec{u}$ y sea v un punto fuera de ella (figura 1.5). Queremos encontrar el punto p en la recta más cercano a v.

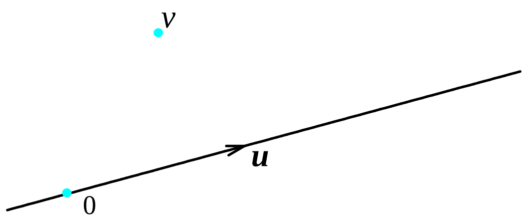

Figure 1.4

Por una parte uno tiene del teorema de Pitágoras que el punto p en la recta generada por $\vec{u}$ y que está más cercano a v debe estar al pie de la perpendicular a nuestra recta y que pasa por v (ver figura 1.5)

En efecto, si q es cualquier ortro punto en la recta generada por el vector $\vec{u}$, entonces, p, q y v son los vértices de un triángulo rectángulo (ver figura 1.5) cuya hipotenusa es de longitud igual a la distancia de v a q y como la hipotenusa es mayor que cualquiera de las catetos se tiene que la distancia de v a p es menor que la distancia de v a q.

Por otra parte, el punto p que minimiza la distancia de v a la recta generada por $\vec{u}$ deberá ser el extremo final de un vector $\vec{p}$ paralelo a $\vec{u}$. Por lo tanto, si al vector de posición de v lo denotamos por $\vec{v}$, entonces, el vecor $\vec{v} - \vec{p}$ es paralelo al segmento que va de p a v y de la misma longitud.

Así que $\vec{v} - \vec{p}$ debe ser ortogonal a $\vec{u}$ y la distancia más corta de v al subespacio generado por $\vec{u}$ es justamente $\|\vec{v} - \vec{p}\|$.

Si ahora consideramos el caso de un plano Π que pase por el origen y un punto v cualquiera en el espacio.

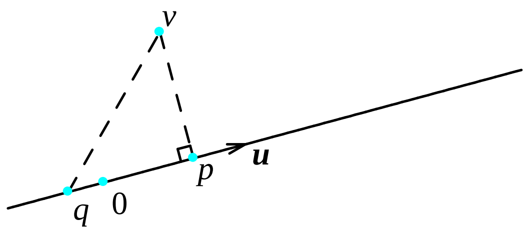

Figure 1.5

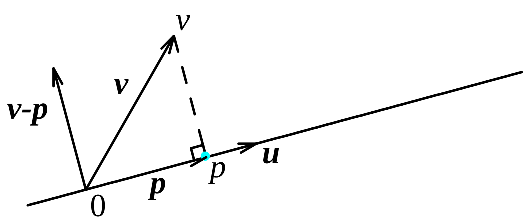

Figure 1.6

Nuevamente, por el teorema de Pitágoras, tenemos que el punto p del plano más cercano al punto v debe ser el pie de la perpendicular al plano que pasa por v (ver figura 1.5). Por supuesto, si denotamos por $\vec{v}$ el vector cuyo extremo final es v y por $\vec{p}$ al vector de posición del punto p, entonces, la distancia más corta del punto v al plano Π es justamente $\|\vec{v}-\vec{p}\|$. Notemos además que el vector $\vec{v}-\vec{p}$ es perpendicular a cualquier vector $\vec{u}$ en el plano Π.

Para el caso general en el que se tiene un elelmento v de un espacio vectorial real con producto interior $\left(V, \langle\cdot,\cdot\rangle\right)$ y un subespacio E es posible demostrar que, en efecto, existe un único $p \in E$ tal que $\|v-p\|_2$ es la distancia más corta de v al subespacio E.

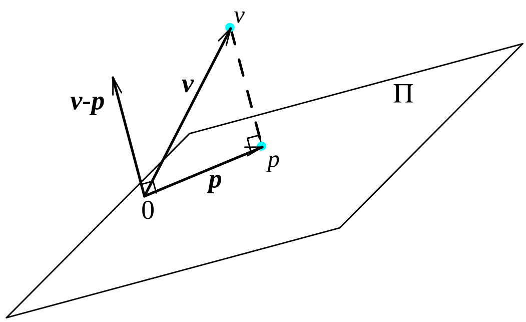

Figure 1.7

TEOREMA 1.5.1 (Teorema de la Proyección). *Sea* $\left(V, \langle\cdot,\cdot\rangle\right)$ *un espacio vectorial de dimension finita con producto interior, sea y sean* E *un subespacio de* V. *Entonces, para cada* v *en* V *existe un único elemento* $P_E(v)$ *en el subespacio* E *que minimiza la distncia de* v *a* E, *esto es*

$$\|v - P_E(v)\|_2 = \inf\{\|v - u\|_2 : u \in E\} = \inf_{u \in E} \|v - u\|_2$$

Más aún,

$$\langle v - P_E(v), u\rangle = 0 \ \ \forall u \in E$$

en otros términos, el vector $v - P_E(v)$ *es ortogonla a todo elemento del subespacio* E.

DEMOSTRACIÓN. La mayor dificultad para probar este resultado es que no se cuenta apriori con el concepto de "*la perpendicular al subespacio* E *que pasa por* v" así que uno debe buscar una manera alternativa de mostrar la existencia del elemento $P_E(v)$ que minimiza la distancia de v a E. La demostración de este hecho requiere de algunas ideas de Cálculo avanzado por lo que prevenimos al lector no familiarizado con los conceptos básicos de espacios métricos.

Sea d la distancia más corta de v a E,

$$d = \inf_{u \in E} \|v - u\|_2$$

De las propiedades del infimo sabemos que para cada $n \in \mathbb{N}$ existe u_n en E para el cual se cumpla que $d \leq \|v - u_n\|_2 < d + \frac{1}{n}$ y por lo tanto para la sucesió $\{u_n\}$ se tiene que

$$\lim_{n \to \infty} \|v - u_n\|_2 = d$$

Mostraremos ahora que $\{u_n\}$ es una sucesión de Cauhy en la norma $\|\cdot\|_2$. Puesto que

$$\|u_n - u_m\|_2 = \|(v - u_m) - (v - u_n)\|_2$$

se sigue de la identidad del paralelogramo (prop. 1.4.8-v) que

$$\|(v - u_m) - (v - u_n)\|_2^2 + \|(v - u_m) + (v - u_n)\|_2^2 \\ = 2\|v - u_m\|_2^2 + 2\|v - u_n\|_2^2$$

y por lo tanto

$$\begin{aligned} \|u_n - u_m\|_2^2 &= 2\|v - u_m\|_2^2 + 2\|v - u_n\|_2^2 - \|(v - u_m) + (v - u_n)\|_2^2 \\ &= 2\|v - u_m\|_2^2 + 2\|v - u_n\|_2^2 - \|2v - (u_n + u_m)\|_2^2 \\ &= 2\|v - u_m\|_2^2 + 2\|v - u_n\|_2^2 - 4\left\|v - \tfrac{1}{2}(u_n + u_m)\right\|_2^2 \end{aligned}$$

Como $\frac{1}{2}(u_n + u_m) \in E$ se tiene que $-4\left\|v - \frac{1}{2}(u_n + u_m)\right\|_2^2 \leq -4d^2$ en consecuencia

$$\|u_n - u_m\|_2^2 \leq 2\|v - u_m\|_2^2 + 2\|v - u_n\|_2^2 - 4d^2$$

y tomando el límite cuando n y m tienden a infinito obtenemos que

$$\lim_{n,m\to\infty} \|u_n - u_m\|_2^2 \leq 2d^2 + 2d^2 - 4d^2 = 0$$

Esto muestra que la sucesión $\{u_n\}$ es de Cauchy y puesto que V es de dimensión finita se sigue que la sucesión $\{u_n\}$ converge a un elemento $P_E(v)$ de E y

$$\|v - P_E(v)\|_2^2 = \lim_{n\to\infty} \|v - u_n\|_2^2 = \inf_{u\in E} \|v - u\|_2^2$$

Esto muestra la existencia del punto $P_E(v)$ que minimiza la distancia de v al subespacio E.

Para demostrar que $P_E(v)$ es único procedemos por contradicción. Suponga que $q \in E$ cumple la condición

$$\|v - q\|_2 = \inf_{u\in E} \|v - u\|_2 = d$$

Entonces, usando nuevamente la identidad del paralelogramo se tiene que

$$\begin{aligned} \|P_E(v) - q\|_2^2 &= 2\|v - P_E(v)\|_2^2 + 2\|v - q\|_2^2 - 4\left\|v - \tfrac{1}{2}(P_E(v) + q)\right\|_2^2 \\ &= 4d^2 - 4\left\|v - \tfrac{1}{2}(P_E(v) + q)\right\|_2^2 \end{aligned}$$

pero $P_E(v) + q \in E$ por lo tanto $-4\left\|v - \frac{1}{2}(P_E(v) + q)\right\|_2^2 \leq -4d^2$ lo que implica que

$$\|P_E(v) - q\|_2^2 \leq 4d^2 - 4d^2 = 0$$

Resta ver que el vector $v - P_E(v)$ es ortogonal a E. Para ello observemos que como $P_E(v) \pm w \in E$ para cualquier $w \in E$

$$\begin{aligned} \|v - P_E(v)\|_2^2 &\leq \|v - (P_E(v) \pm w)\|_2^2 \\ &= \|(v - P_E(v)) \mp w\|_2^2 \\ &= \|v - P_E(v)\|_2^2 \mp 2\langle v - P_E(v), w\rangle + \|w\|_2^2 \end{aligned}$$

por lo tanto

$$2\left|\langle v - P_E(v), w\rangle\right| \leq \|w\|_2^2 \quad \forall w \in E$$

Así, para cualquier $\lambda \in \mathbb{R}$ y para cualquier $u \in E$ se cumple que

$$2\left|\langle v - P_E(v), \lambda u\rangle\right| \leq \|\lambda u\|_2^2$$

y, en cosecuencia,

$$2\left|\langle v - P_E(v), u\rangle\right| \leq |\lambda|\,\|u\|_2^2 \quad \forall \lambda \in \mathbb{R} \quad \text{y} \quad \forall u \in E$$

En particula para $\lambda = 0$ obtenemos que

$$\left|\langle v - P_E(v), u\rangle\right| = 0$$

□

DEFINICIÓN 1.5.2. Sea $\left(V, \langle\cdot,\cdot\rangle\right)$ un espacio vectorial real con producto interior y sea E un subespacio de V. Para cada vector v en V definimos **la proyección ortogonal de** v **sobre** E como el vector $P_E(v)$ en V que minimiza la distancia de v al subespacio E.

OBSERVACIÓN 2. El teorema de la proyección nos permite darle sentido a la epresión "*la perpendicular al subespacio E que pasa por* v". En efecto, de acuerdo con el teorema el vector $v - P_E(v)$ es ortogonal al subespacio E y por ende nos da la dirección de la perpendicular a E que pasa por v.

1.6. Bases ortogonales y ortonormales

Una de las herramiantas más importantes de los espacios con producto interior son los conjuntos de vectores mutuamente perpendiculares u ortogonales y, en particular, el de bases de vectores mutuamente perpendiculares. En esta sección preisaremos el concepto de "*base de vecores mutuamente perpendiculares*", discutiremos algunas de sus propiedades y desarrollaremos un método para construir dichas bases.

DEFINICIÓN 1.6.1. Dado un espacio vectorial con producto interior $\left(V, \langle\cdot,\cdot\rangle\right)$. Diremos que un conjunto de vectores $\{u_1, ..., u_m\}$ en V es un **conjunto perpendicular u ortogonal** si cumplen la siguiente condición

$$\langle u_k, u_j\rangle = 0 \text{ cuando } u_k \neq u_j$$

cum Una base $\{u_1, ..., u_n\}$ de V es llamada una **bae ortogonal** si es un conjunto ortogonla y una base de V. Si además los elementos de dicha base tienen norma 1, $\|u_k\|_2 = 1$,diremos que la **base es ortonormal**.

Una primera propiedad de los conjuntos ortogonales es que si ningún elemento es cero, los vectores son linealmente independientes. En efecto, si $\{u_1, \dots, u_m\}$ es un conjunto ortogonal con $u_k \neq 0$ para $k = 1, \dots, m$ y

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m = 0$$

entonces, al hacer el producto interior con u_1 obtenemos que

$$\langle \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m, u_1 \rangle = \langle 0, u_1 \rangle = 0$$

y como

$$\begin{aligned} &\langle \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m, u_1 \rangle \\ &\qquad = \alpha_1 \langle u_1, u_1 \rangle + \alpha_2 \langle u_2, u_1 \rangle + \cdots + \alpha_m \langle u_m, u_1 \rangle \end{aligned}$$

se sigue que

$$\alpha_1 \langle u_1, u_1 \rangle + \alpha_2 \langle u_2, u_1 \rangle + \cdots + \alpha_m \langle u_m, u_1 \rangle = 0$$

pero $\langle u_k, u_1 \rangle = 0$ para $k \neq 1$, por lo tanto

$$\alpha_1 \langle u_1, u_1 \rangle = 0$$

y puesto que $u_1 \neq 0$, necesariamente $\alpha_1 = 0$. Si procedemos de modo similar con $u_2, \dots, u_m$ se concluye que $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$ por lo que el conjunto es linealmente independiente. Resumiendo

TEOREMA 1.6.2. *Sea $(V, \langle \cdot, \cdot \rangle)$ un espaci vectorial con producto interior. Entonces todo conjunto ortogonal $\{u_1, \dots, u_m\}$ que no contenga al cero es linelamente independiente.*

Una característica fundamental de las bases ortogonales es la siguiente:

TEOREMA 1.6.3. *Si $\{u_1, ..., u_n\}$ es una base ortogonal para un espacio con producto interior $(V, \langle \cdot, \cdot \rangle)$, entonces, para cualquier vector v en V*

$$v = \frac{\langle v, u_1 \rangle}{\|u_1\|_2^2} u_1 + \cdots + \frac{\langle v, u_n \rangle}{\|u_n\|_2^2} u_n = \sum_{k=1}^{n} \frac{\langle v, u_k \rangle}{\|u_k\|_2^2} u_k \tag{1.6}$$

Si $\{u_1, ..., u_n\}$ es una base ortonormal, entonces

$$v = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_n \rangle u_n = \sum_{k=1}^{n} \langle v, u_k \rangle u_k \tag{1.7}$$

*A las expresiones de la forma (1.6) y (1.7) se les denomina **series o representaciones del tipo de Fourier.***

DEMOSTRACIÓN. Sea $\{u_1, ..., u_n\}$ es una base ortogonal y v es un vector cualquiera de V. Puesto que $\{u_1, ..., u_n\}$ es una base existen coeficientes $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$v = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$$

Al hacer el producto interior con u_1 nos queda que

$$\begin{aligned} \langle v, u_1 \rangle &= \langle \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m, u_1 \rangle \\ &= \alpha_1 \langle u_1, u_1 \rangle + \alpha_2 \langle u_2, u_1 \rangle + \cdots + \alpha_n \langle u_n, u_1 \rangle \end{aligned}$$

y como $\langle u_k, u_1 \rangle = 0$ para $k \neq 1$, concluimos que

$$\langle v, u_1 \rangle = \alpha_1 \langle u_1, u_1 \rangle = \alpha_1 \|u_1\|_2^2$$

por lo tanto

$$\alpha_1 = \frac{\langle v, u_1 \rangle}{\|u_1\|_2^2}$$

repitiendo el argumento con $u_2, \ldots, u_n$ se obtiene que

$$\alpha_1 = \frac{\langle v, u_1 \rangle}{\|u_1\|_2^2}, \alpha_2 = \frac{\langle v, u_2 \rangle}{\|u_2\|_2^2}, \ldots, \alpha_n = \frac{\langle v, u_n \rangle}{\|u_n\|_2^2}$$

Así

$$v = \frac{\langle v, u_1 \rangle}{\|u_1\|_2^2} u_1 + \cdots + \frac{\langle v, u_n \rangle}{\|u_n\|_2^2} u_n = \sum_{k=1}^{n} \frac{\langle v, u_k \rangle}{\|u_k\|_2^2} u_k$$

Si además la base es ortonormal, entonces $\|u_k\|_2^2 = 1$ para $k = 1, \ldots, n$ por lo que la expresión anterior se reduce a la forma

$$v = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_n \rangle u_n = \sum_{k=1}^{n} \langle v, u_k \rangle u_k \qquad \square$$

Otra ventaja de las bases ortonormales es que la magnitud de un vector se puede calcular de modo similar a como se hace con los vectores en $\mathbb{R}^n$, esto es, como la raíz de la suma de los cuadrados de las componentes:

TEOREMA 1.6.4. *Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interior y sea $\{u_1, ..., u_n\}$ una base ortonormal, entonces, para cualquier vector $v \in V$ se tiene que*

$$\|v\|_2 = \sqrt{\langle v, u_1 \rangle^2 + \cdots + \langle v, u_n \rangle^2} = \sqrt{\sum_{k=1}^{n} \langle v, u_k \rangle^2}$$

Si la base es solamente ortogonal, entonces

$$\|v\|_2 = \sqrt{\frac{\langle v, u_1 \rangle^2}{\|u_1\|_2^2} + \cdots + \frac{\langle v, u_n \rangle^2}{\|u_n\|_2^2}} = \sqrt{\sum_{k=1}^{n} \frac{\langle v, u_k \rangle^2}{\|u_k\|_2^2}}$$

DEMOSTRACIÓN. Sabemos que $\|v\|_2^2 = \langle v, v\rangle$ y que $v = \langle v, u_1\rangle\, u_1 + \cdots + \langle v, u_n\rangle\, u_n$ así que

$$\|v\|_2^2 = \left\langle \sum_{k=1}^{n} \langle v, u_k\rangle\, u_k, \sum_{k=1}^{n} \langle v, u_k\rangle\, u_k \right\rangle$$

Si se usan las propiedades i) y ii) de la definición de producto interior (Definición 1.4.1) obtenemos que

$$\begin{aligned}\|v\|_2^2 &= \langle v, u_1\rangle \left\langle \sum_{k=1}^{n} \langle v, u_k\rangle\, u_k, u_1 \right\rangle + \cdots + \langle v, u_n\rangle \left\langle \sum_{k=1}^{n} \langle v, u_k\rangle\, u_k, u_n \right\rangle \\ &= \langle v, u_1\rangle \sum_{k=1}^{n} \langle v, u_k\rangle \langle u_k, u_1\rangle + \cdots + \langle v, u_n\rangle \sum_{k=1}^{n} \langle v, u_k\rangle \langle u_k, u_n\rangle\end{aligned}$$

Pero $\langle u_k, u_j\rangle = 0$ excepto cuando $k = j$ por lo que esta expresión se reduce a la forma

$$\begin{aligned}\|v\|_2^2 &= \langle v, u_1\rangle^2 \langle u_1, u_1\rangle + \cdots + \langle v, u_n\rangle^2 \langle u_n, u_n\rangle \\ &= \langle v, u_1\rangle^2 \|u_1\|_2^2 + \cdots + \langle v, u_n\rangle^2 \|u_n\|_2^2\end{aligned}$$

Finalmente como $\|u_k\|_2 = 1$ concluimos que

$$\|v\|_2^2 = \langle v, u_1\rangle^2 + \cdots + \langle v, u_n\rangle^2$$

y el rsultado se sigue.

El caso de una base ortogonal es similar a este y se deja como ejercicio. $\square$

1.7. El método de Gram-Schmidt

Como se vió en la sección anterior, las bases ortogonales y ortonormales ofrecen grandes ventajas, por lo que resulta relevante poder construir conjutos y bases ortogonales. En esta sección discutiremos un método que nos permite construir un conjunto ortogonal a partir de un conjunto de vectores linelamente independientes.

Veamos primero el caso en el se tienen dos vectores u y v linealmente independietes. Como se vió anterirmente una maera de construir vectores mutuamente perpendiculare es considerando, digamos la proyección ortogonal de v sobre el subespacio generado por u

En efecto, sabemos del Teorema 1.5.1 que si p es el vector proyección de v sobre u, entonces $v - p$ es perpendicular a u. Así el par de vectores u y $v - p$ forman un conjunto ortogonal.

Resta ver como calcular, en este caso, el vector p en términos de u y v. Para ello, obsevemos que como p esta en el subespacio generado

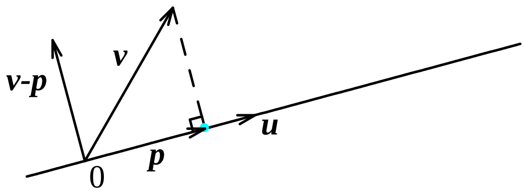

Figure 1.8

por u, p debe ser de la forma λu para algún valor de λ y por tanto $v - p = v - \lambda u$ es perpendicular a u, esto es

$$\langle v - \lambda u, u\rangle = \langle v - p, u\rangle = 0$$

De las propiedades del producto interior obtenemos que

$$\langle v - \lambda u, u\rangle = \langle v, u\rangle - \lambda \langle u, u\rangle = 0$$

y al despejar λ nos queda

$$\lambda = \langle v, u\rangle \;/\; \|u\|_2^2$$

Así, el vector proyección está dado por

$$p = \frac{\langle v, u\rangle}{\|u\|_2^2} u$$

y el conjunto $\left\{u, v - \frac{\langle v,u\rangle}{\|u\|_2^2} u\right\}$ es un conjunto ortogonal.

Observemos además que u y v generan el mismo subespacio que u y $v - \frac{\langle v,u\rangle}{\|u\|_2^2} u$. En efecto, si $w = \alpha u + \beta v$, entonces, $w = \left(\alpha + \beta \frac{\langle v,u\rangle}{\|u\|_2^2}\right) u + \beta \left(v - \frac{\langle v,u\rangle}{\|u\|_2^2} u\right)$ y por supuesto cualquier combinación lineal de u y $v - \frac{\langle v,u\rangle}{\|u\|_2^2} u$ se puede escribir en términos de u y v.

Notación 2. Resulta conveniente para la discución que sigue usar la expresión $[[w_1, \ldots, w_k]]$ para denotar al subespacio generado por los vectores $w_1, \ldots, w_k$. Esto es,

$$[[w_1, \ldots, w_k]] = \{w \in V : w = \alpha_1 w_1 + \cdots + \alpha_k w_k \text{ con } \alpha_1, \ldots, \alpha_k \in \mathbb{R}\}$$

Consideremos ahora el caso de tres vectores linealmente independientes $\{v_1, v_2, v_3\}$. Tomemos primero dos vectores, digamos, v_1, v_2 y construyamos los vectores

$$u_1 = v_1$$

y

$$u_2 = v_2 - \frac{\langle v_2, v_1\rangle}{\|v_1\|_2^2} v_1 = v_2 - \frac{\langle v_2, u_1\rangle}{\|u_1\|_2^2} u_1$$

Por lo visto previamente, sabemos que u_1 y u_2 son ortogonales y generan el mismo subespacio que v_1 y v_2, esto es, $[[v_1, v_2]] = [[u_1, u_2]]$. Para construir un tercer vector perpendicular, considermos la proyección de v_3 sobre el subespacio $[[v_1, v_2]]$, esto es, $P_{[[v_1,v_2]]}(v_3)$, entonces, puesto que $v_3 - P_{[[v_1,v_2]]}(v_3)$ es ortogonal a todo vector en $[[v_1, v_2]]$ y u_1, u_2 están en $[[v_1, v_2]]$ se tiene que $v_3 - P_{[[v_1,v_2]]}(v_3)$ es perpendicular a u_1 y u_2 lo que muestra que

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2 - \frac{\langle v_2, u_1\rangle}{\|u_1\|_2^2} u_1 \\ u_3 &= v_3 - P_{[[v_1,v_2]]}(v_3) \end{aligned}$$

es una base ortogonal de $[[v_1, v_2, v_3]]$. Más aún, usando la base ortogonal $\{u_1, u_2\}$ de $[[v_1, v_2]]$ podemos calcular $P_{[[v_1,v_2]]}(v_3)$. En efecto, sabemos que

$$P_{[[v_1,v_2]]}(v_3) = \frac{\langle P_{[[v_1,v_2]]}(v_3), u_1\rangle}{\|u_1\|_2^2} u_1 + \frac{\langle P_{[[v_1,v_2]]}(v_3), u_2\rangle}{\|u_2\|_2^2} u_2$$

por otra parte, como $v_3 - P_{[[v_1,v_2]]}(v_3) \perp u_1, u_2$

$$\langle v_3 - P_{[[v_1,v_2]]}(v_3), u_1\rangle = 0 \text{ y } \langle v_3 - P_{[[v_1,v_2]]}(v_3), u_2\rangle = 0$$

de donde se obtiene que

$$\langle v_3, u_1\rangle = \langle P_{[[v_1,v_2]]}(v_3), u_1\rangle \text{ y } \langle v_3, u_2\rangle = \langle P_{[[v_1,v_2]]}(v_3), u_2\rangle = 0$$

y por ende

$$P_{[[v_1,v_2]]}(v_3) = \frac{\langle v_3, u_1\rangle}{\|u_1\|_2^2} u_1 + \frac{\langle v_3, u_2\rangle}{\|u_2\|_2^2} u_2$$

En particular

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2 - \frac{\langle v_2, u_1\rangle}{\|u_1\|_2^2} u_1 \\ u_3 &= v_3 - \frac{\langle v_3, u_1\rangle}{\|u_1\|_2^2} u_1 - \frac{\langle v_2, u_2\rangle}{\|u_2\|_2^2} u_2 \end{aligned}$$

es una base ortogonal para $[[v_1, v_2, v_3]]$

TEOREMA 1.7.1 (Gram-Schmidt). *Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio vectorial con producto interior de dimensión finita, entonces, si $\{v_1, ..., v_k\}$ es un conjunto de vectores linealmente independientes, los vectores $\{u_1, ..., u_k\}$ dados por la fórmula*

$$\begin{aligned} u_1 &= v_1 \\ u_j &= v_j - \sum_{i=1}^{j-1} \frac{\langle v_j, u_i\rangle}{\|u_i\|^2} u_i, \quad j = 2, \ldots, k \end{aligned}$$

forman un conjunto de vectores mutuamente ortogonales que generan el mismo subespacio que el conjunto $\{v_1, ..., v_k\}$. En particular si $\{v_1, ..., v_n\}$ es una base de V, entonces, $\{u_1, ..., u_n\}$ es una base de vectores mutuamente ortogonales para V.

DEMOSTRACIÓN. La prueba es por inducción. Para $k = 1$ no hay nada que demostrar. Para $k = 2$ se tiene que

$$u_1 = v_1 \text{ y } u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{\|u_1\|^2} u_1$$

y por lo tanto

$$\begin{aligned} \langle u_2, u_1\rangle &= \langle v_2, u_1\rangle - \frac{\langle v_2, u_1\rangle}{\|u_1\|^2} \langle u_1, u_1\rangle \\ &= \langle v_2, u_1\rangle - \frac{\langle v_2, u_1\rangle}{\|u_1\|^2} \|u_1\|^2 = 0 \end{aligned}$$

Además como

$$v_1 = u_1 \text{ y } v_2 = u_2 + \frac{\langle v_2, u_1\rangle}{\|u_1\|^2} u_1$$

$[[v_1, v_2]] = [[u_1, u_2]]$.

Supongamos que el resultado es válido para $k = m$ y sea $\{v_1, \ldots, v_m, v_{m+1}\}$ un conjunto linealmente independiente y sean

$$u_1 = v_1 \text{ y } u_j = v_j - \sum_{i=1}^{j-1} \frac{\langle v_j, u_i\rangle}{\|u_i\|^2} u_i, j = 2, \ldots, m+1$$

Por hipótesis de inducción $\{u_1, ..., u_m\}$ son mutuamente perpendiculares y generan el mismo subespacio que $\{v_1, ..., v_m\}$. Sea

$$u_{m+1} = v_{m+1} - \sum_{i=1}^{m} \frac{\langle v_{m+1}, u_i\rangle}{\|u_i\|^2} u_i$$

entonces para $j = 1, \ldots, m$ se tiene que

$$\langle u_{m+1}, u_j \rangle = \langle v_{m+1}, u_j \rangle - \sum_{i=1}^{m} \frac{\langle v_{m+1}, u_i \rangle}{\|u_i\|^2} \langle u_i, u_j \rangle$$

Como $\langle u_i, u_j \rangle = 0$ si $i \neq 0$

$$\begin{aligned} \langle u_{m+1}, u_j \rangle &= \langle v_{m+1}, u_j \rangle - \frac{\langle v_{m+1}, u_j \rangle}{\|u_j\|^2} \langle u_j, u_j \rangle \\ &= \langle v_{m+1}, u_j \rangle - \frac{\langle v_{m+1}, u_j \rangle}{\|u_j\|^2} \|u_j\|^2 = 0 \end{aligned}$$

por lo tanto u_{m+1} es ortogonal a u_j para $j = 1, \ldots, m$, esto es, $\{u_1, ..., u_m, u_{m+1}\}$ es un conjunto ortogonal y como

$$v_{m+1} = u_{m+1} + \sum_{i=1}^{m} \frac{\langle v_{m+1}, u_i \rangle}{\|u_i\|^2} u_i$$

se tiene que $[[v_1, \ldots, v_m, v_{m+1}]] = [[u_1, \ldots, u_m, v_{m+1}]] = [[u_1, \ldots, u_m, u_{m+1}]]$. □

1.8. Problemas

Los problemas 1 a 8 están dedicados a reformular algunos conceptos geométricos en términos del producto interior.

1. ***a***) Si u, v y w son tres vectores linealmente depaendientes y de norma uno, se dice que v está entre u y w si $\langle u, w \rangle - \langle u, v \rangle \langle w, v \rangle \leq 0$. ¿Podría dar un argumento geométrico para justificar este criterio? (Sugerencia: Considere los vectores $u - \langle u, v \rangle v$ y $w - \langle w, v \rangle v$ y recuerde que los vectores son de norma uno).

   b) Como se vió, el coseno del ángulo entre dos elementos no nulos de un espacio vectorial con producto interior se puede definir como

   $$\cos\theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}. \quad 0 \leq \theta \leq \pi$$

   En consecuencia, para definir el seno del ángulo entre dos vectores no nulos usando solo el producto interior se puede usar la relación $\operatorname{sen}\theta = \sqrt{1 - \cos^2\theta}$, esto es

   $$\operatorname{sen}\theta = \sqrt{1 - \frac{\langle u, v \rangle^2}{\|u\|^2 \|v\|^2}}$$

   Demuestre usando solo estas definiciones y los propiedades del producto interior que se cumplen las siguientes identidades trigonométricas

   $$\cos(\alpha \pm \beta) = \cos\alpha \cos\beta \mp \operatorname{sen}\alpha \operatorname{sen}\beta$$

(Sugerencia: Considere tres vectores u, v y w de norma uno, con v entre u y w y exprese w en términos de la base ortogonal $\{v, u - \langle u, v \rangle v\}$ y use el producto interior)

2. Muestre que la suma de los ángulos internos de un triángulo es π. (Sugerencia: $\alpha + \beta + \gamma = \pi$ si y solo si $\alpha + \beta = \pi - \gamma$ y recuerde que $\cos(\pi - \gamma) = -\cos\gamma$).

3. Demuestre que, en un triángulo isósceles, las medianas que van a los lados iguales son iguales y, resíprocamente, si en un triángulo dos medianas son iguales, el triángulo es isósceles.

4. Demuestre que la suma de los cuadrados de las tres medianas de un triangulo es $3/4$ de la suma de los cuadrados de los lados.

5. Demuestre que en un paralelogramo la suma de los cuadrados de las diagonales es igual a la suma del cuadrado de los lados.

6. ***a***) Muestre que dada la recta $ax + by + c = 0$ el vector $\bar{n} = (a, b)$ es perpendicular a cualquier vector en la dirección de la recta y exprese la ecuación de la recta en términos de $\bar{n}$ y un punto en la recta usando el producto punto.

b) Exprese la ecuación de un plano en $\mathbb{R}^3$ en términos del producto punto usando un vector ortogonal al plano y un punto en este.

c) Sea $\bar{n} \neq \bar{0}$ un vector en $\mathbb{R}^4$ y P_0 un punto de $\mathbb{R}^4$. ¿Como interpretaría geometricamente al conjunto solución de la ecuación $\bar{n} \cdot (\bar{P} - \bar{P}_0) = 0$? (Al conjunto soución de esta ecuación se le denomina un hiperplano en $\mathbb{R}^4$)

d) ¿Cuál sería su definición de hiperplano en un espacio vectorial con producto interior?

e) Muestre un hiperplano que pasa por el vector 0 en un espacio vectorial $\left(V, \langle \cdot, \cdot \rangle\right)$ de dimensión n es un subespacio de dimensión $n - 1$.

7. ***a***) La fórmula de la distancia de un punto (ξ, η) a una recta $ax + by + c = 0$ es

$$\left| \frac{a\xi + b\eta + c}{\sqrt{a^2 + b^2}} \right|$$

Reescriba esta expresión en términos del producto interior y de una interpretación geometrica de la fórmula.

b) Reescriba la fórmula de la distncia de un punto a un plano en en términos del producto punto.

c) ¿Cuál sería la fórmula de la distancia de un punto a un hiperplano en un espacio vectorial con producto interior?

8. ***a***) Muestre que un círculo con centro en P_0 y radio r en el plano se puede describir como el conjunto de puntos P tales que $\|\bar{P} - \bar{P}_0\| = r$ y que una esfera en $\mathbb{R}^3$ está dada por los puntos $P \in \mathbb{R}^3$ tales que $\|\bar{P} - \bar{P}_0\| = r$. ¿Cuál sería su definición de una "hiperesfera" en un espacio vectorial con porducto interior?

b) Muestre que el disco en el plano con centro en P_0 y radio r está descrito por los puntos $P \in \mathbb{R}^2$ tales que $\|\bar{P} - \bar{P}_0\| < r$ y que una bola en $\mathbb{R}^3$ con centro en P_0 y radio r es el conjunto de puntos $P \in \mathbb{R}^3$ que cumplen la condición $\|\bar{P} - \bar{P}_0\| < r$. ¿Cuál sería su definición de una "hiperbola" en un espacio vectorial con porducto interior?

9. En los ejemplos 1.1, 1.2, 1.3 y 1.4 muestre que los productos ahí definidos son, en efecto productos interiores.

10. En el ejemplo 1.1 muestre que $(1,0)$ no tiene longitud 1 pero $(\frac{1}{2},0)$ si lo es. ¿Cual es el ángulo entre $(1,0)$ y $(1,1)$ con este producto interior?

11. ***a***) En $P_4[x]$ con el producto interior del ejemplo 1.2 use Gram-Schmidt para encontrar una base ortonormal a partir de la base $\{1, x, x^2, x^3, x^4\}$.

b) En $P_4[x]$ con el producto interior del ejemplo 1.4 use Gram-Schmidt para encontrar una base ortonormal a partir de la base $\left\{1, x, x^2, x^3, x^4\right\}$.

c) En $P_4[\cos x]$ con el producto interior del ejemplo 1.3 use Gram-Schmidt para encontrar una base ortonormal a partir de la base $\left\{1, \cos x, \cos^2 x, \cos^3 x, \cos^4 x\right\}$.

12. ***a***) Exprese el vector $x^3 + x^2 \in P_4[x]$ en términos de la base ortogonal obtenida en el inciso b) del problema anterior.

b) Muestre que $\operatorname{sen}^4 x$ está en el espacio vectorial $P_4[\cos x]$ del ejemplo 1.3 y exprese este vector en términos de la base obtenida en el inciso c) del problema anterior.

13. Demuestre todos los resultados que se dejan como ejercicio a lo largo de este capítulo.

Capítulo 2

Operadores Adjuntos, de Proyección y Ortogonales

En este capitulo haremos un recuento de algunas de las propiedades de ciertas clases de operadores que frecuentemente aparecen en diferentes problemas de álgebra lineal y análisis.

2.1. Operadores Lineales y Matrices

En esta sección daremos un breve repaso de operadores lineales y de la manera en que se puede asociar una matriz a un operador lineal, en espacios vectoriales de dimensión finita.

Dados espacios vectoriales U y V, diremos que una función $T\colon U \to V$ es un *operador lineal* si

$$T(\alpha\vec{u}_1 + \beta\vec{u}_2) = \alpha T\vec{u}_1 + \beta T\vec{u}_2 \quad \forall \vec{u}_1, \vec{u}_2 \in V \; \forall \alpha, \beta \in \mathbb{R}$$

Con cada operador lineal uno tiene definidos dos conjuntos el *núcleo* o *kernel* de T y el *rango o imagen* de T. El primero se define como el conjunto

$$\ker T = \{\vec{u} \in U \mid T\vec{u} = 0\}$$

y el segundo esta dado por el conjunto

$$\Im \{\vec{v} \in V \mid \exists \vec{u} \in U \text{ tal que } T\vec{u} = \vec{v}\}$$

Una de las propiedades básicas de los operadores lineales, es que tanto su kernel como su rango son subespacios de U y V, respectivamente.

Sean U y V espacios de dimensión finita, y sea $T : U \to V$ un operador lineal. Entonces, dadas una base de U y una base de V, uno puede asociar a T una matriz mediante el siguiente procedimiento:

Denotemos por $\mathcal{U} = \{\vec{u}_1, \dots, \vec{u}_n\}$ a la base de U y por $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_m\}$ a la base de V. Sabemos que $T\vec{u}_j$ se puede representar de marera única en la base $\mathcal{V}$ como

$$T\vec{u}_j = a_{1j}\vec{v}_1 + \cdots + a_{mj}\vec{v}_m$$

La matriz asociada a T con respecto a las bases $\mathcal{U}$ y $\mathcal{V}$ es la matriz, A, cuyas columnas son las componentes de los vectores $T\vec{u}_j$ en la base $\mathcal{V}$:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

La relación entre esta matriz y el operador lineal T es la siguiente: Si $\vec{x}$ se expresa en la base $\mathcal{U}$ como

$$\vec{x} = x_1\vec{u}_1 + \cdots + x_n\vec{u}_n$$

y $\vec{y} = T\vec{x}$ está dado por $\vec{y} = y_1\vec{v} + \cdots + y_m\vec{v}_m$ en la base $\mathcal{V}$, entonces,

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

En forma más sintética $A\vec{x} = \vec{y}$.

También se puede establecrer el recíproco, si $\mathcal{U} = \{\vec{u}_1, \dots, \vec{u}_n\}$ es base de U y $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_m\}$ es base de V, entonces, a cada matriz le corresponde un único operador lineal $T\colon U \to V$.

Observación 3. Es importante tener en mente que la matriz asociada a un operador lineal, depende tanto de la base que se elija en el dominio, U, como de la que se elija en el contradominio, V. Este hecho plantea el siguiente problema: Dada una transformación lineal $T\colon U \to V$ ¿Cuales son las bases de U y V en las que T tiene una representación matricial lo más sencilla posible? Tratar de responder a esta clase de pregunta es uno de los principales objetivos de este curso.

Notación 3. Cuando trabajemos con matrices, usaremos la notación $A = (a_{ij})$ para indicar que la matriz A tiene como entrada en el renglón i, columna j el valor a_{ij}. Esto es, el primer subíndice indica el renglón en el que está colocado el valor a_{ij} y el segundo subíndice la columna que le corresponde. Si $U = V$ nosotrros hablaremos tan sólo de la matriz asociada en la base $\mathcal{U}$, cuando la base U se toma tanto en el dominio como en el contradominio. Denotaremos por $I_V\colon V \to V$ al operador identidad $I_V\vec{x} = \vec{x}$ y a la corespondiente matriz identidad, y cuando no haya posibilidad de confusión suprimiremos el subídice.

Cuando se trabaja con espacios de dimensión finita, una transformación lineal $T\colon V \to V$ inyectiva siempre tiene una inversa $T^{-1}\colon V \to V$, tal que $TT^{-1} = I$ y $T^{-1}T = I$. Sin embargo, en espacios más generales, estos dos conceptos no son del todo equivalentes. Por esta razón nosotros usaremos el término *no singular* para indicar que nuestro operador es inyectivo y el término invertible para indicar que T tiene una inversa T^{-1} tal que $TT^{-1} = I$ y $T^{-1}T = I$. Recuerdese que en le caso en que $T\colon U \to V$ con $U \neq V$, existe una variante en esta definición de invertibilidad; el operador inverso T^{-1} debe cumplir las condiciones $TT^{-1} = I_V$ y $T^{-1}T = I_U$. Si esto ocurre diremos que T es invertible. Por supuesto, el operador inverso es único.

Es importante tener presente que, en espacios de dimensión finita, uno siempre puede asociar a cada operador $T\colon V \to V$ su *determinante*, $\det T$, el cual está dado por el determinante de la matriz asociada a T en cualquier base (se puede mostrar que el valor del determinante es independiente de la base que se elija para representar a T como matriz).

Algunas de los critérios para determinar cuando un operador lineal en diemensión finita es invertible, son los siguientes:

- El detereminante nos da un criterio para saber si un operador lineal es invertible: *T es invertible si y sólo si* $\det T \neq 0$
- Otro criterio útil para determinar cuando un operador $T : V \to V$ es no singular, es el del kernel: T es no singular si y sólo si $\ker T = \{0\}$.
- Estos dos criterios son equivalentes cuando se trabaja en espacios de dimensión finita. Su contra parte también es muy usada: $\ker T \neq \{0\}$ si y sólo si $\det T = 0$.

Terminamos esta sección señalando un hecho que resulta de utilidad en problemas concretos.

Cuando uno cuenta con espacios vectoriales con poducto interior y las base $\mathcal{U}$ y $\mathcal{V}$ son ortonormales, el cálculo de la matriz asociada a un operador lineal $T : U \to V$ se reduce a calcular los productos $\langle T\vec{u}_j, \vec{v}_i\rangle$, $i = 1, ..., m$, $j = 1, ..., n$. En efecto, uno puede mostrar que si $a_{ij} = \langle T\vec{u}_j, \vec{v}_i\rangle$, entonces, la matriz asociada a T es la matriz $A = (a_{ij})$. La verificación de este resultado se deja como ejercicio.

2.2. El Operador Adjunto

Consideremos ahora un espacio vectorial de dimensión finita con producto interior $(V, \langle\cdot,\cdot\rangle)$. En este caso, con cada operador lineal $T : V \to V$, el producto interior permite asociar, de manera natural, un segundo operador lineal $T^* : V \to V$ el cual se define, en principio, mediante la fórmula

$$\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, T^*\vec{v}\rangle , \forall \vec{u}, \vec{v} \in V$$

Esta fórmula no garantiza *a priori* que T^* defina una función y mucho menos el que T^* sea un operador lineal. Para que este sea el caso, es necesario establecer la siguiente condición:

Para cada $\vec{v}$ en V existe un único $\vec{w}$ en V tal que, $\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, \vec{w}\rangle$ para todo $\vec{u}$ en V.

Si esta propiedad se cumple, la relación que a cada $\vec{v}$ le asigna el vector $\vec{w}$, define una función T^* de V en V y uno puede demostrar, usando las propiedades de producto interior, que T^* es lineal y que $\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, T^*\vec{v}\rangle$ para todo $\vec{u}$ y $\vec{v}$ en V.

LEMA. *Sea $(V, \langle\cdot,\cdot\rangle)$ es un espacio de dimención finita con producto interior. Sea $T\colon V \to V$ un operador lineal, entonces*

i) Para cada v en V esxiste un único w en V tal que $\langle Tu, v\rangle = \langle u, w\rangle$ $\forall u \in V$
ii) La relación $T^*v = w$ define un operador lineal con la propiedad de que

$$\langle Tu, v\rangle = \langle u, T^*v\rangle \quad \forall u, v \in V$$

Además, T^* es el único operador con esta propiedad.

DEMOSTRACIÓN. Sabemos que V posee una base ortonormal, digamos $\{\vec{u}_1, \ldots, \vec{u}_n\}$, por lo tanto, todo $\vec{u}$ en V se puede escribir de la forma $\vec{u} = a_1\vec{u}_1 + \cdots + a_n\vec{u}_n$ con $a_i = \langle \vec{u}, \vec{u}_i\rangle$. Dado $\vec{v}$ en V definimos $\vec{w}$ como el vector $\vec{w} = \sum_{i=1}^{n} \langle T\vec{u}_i, \vec{v}\rangle\, \vec{u}_i$. Entonces, para toda $\vec{u}$ en V se tiene, de la linealidad del producto interior y del hecho de que $\{\vec{u}_1, \ldots, \vec{u}_n\}$ es una base ortonormal

$$\begin{aligned}
\langle \vec{u}, \vec{w}\rangle &= \left\langle \vec{u}, \sum_{i=1}^{n} \langle T\vec{u}_i, \vec{v}\rangle\, \vec{u}_i \right\rangle = \sum_{i=1}^{n} \langle \vec{u}, \langle T\vec{u}_i, \vec{v}\rangle\, \vec{u}_i\rangle \\
&= \sum_{i=1}^{n} \langle T\vec{u}_i, \vec{v}\rangle \langle \vec{u}, \vec{u}_i\rangle = \sum_{i=1}^{n} \langle T\vec{u}_i, \vec{v}\rangle\, a_i \\
&= \sum_{i=1}^{n} \langle a_i T\vec{u}_i, \vec{v}\rangle = \left\langle \sum_{i=1}^{n} a_i T\vec{u}_i, \vec{v} \right\rangle
\end{aligned}$$

Por la linealidad del producto interior y de T, obtenemos que

$$\langle \vec{u}, \vec{w}\rangle = \left\langle \sum_{i=1}^{n} a_i T\vec{u}_i, \vec{v} \right\rangle = \left\langle T\sum_{i=1}^{n} a_i \vec{u}_i, \vec{v} \right\rangle = \langle T\vec{u}, \vec{v}\rangle \;\; \forall \vec{u} \in V$$

Para mostrar la unicidad, supongamos que $\vec{w}'$ también satisface la relación $\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, \vec{w}'\rangle$, $\forall \vec{u} \in V$. Sabemos que $\vec{w}' = \sum_{i=1}^{n} \langle \vec{u}_i, \vec{w}'\rangle \vec{u}_i$, pero $\langle \vec{u}_i, \vec{w}'\rangle = \langle T\vec{u}_i, \vec{v}\rangle$, en consecuencia $\vec{w}' = \sum_{i=1}^{n} \langle T\vec{u}_i, \vec{v}\rangle\, \vec{u}_i = \vec{w}$. (Es este tipo de argumento el que sugirió la definición de $\vec{w}$).

Es claro de *i*) que $T^*\vec{v} = \vec{w}$ define una función y la linealidad se sigue inmediatamente de las propiedades de producto interior y se deja como ejercicio. Resta ver que T^* es el único operador lineal con la propiedad de que

$$\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, T^*\vec{v}\rangle, \forall \vec{u}, \vec{v} \in V$$

Supongamos que S también tiene esa propiedad, entonces, $\langle\vec{u}, S\vec{v}\rangle = \langle\vec{u}, T^*\vec{v}\rangle$ para toda $\vec{v} \in V$ o, lo que es lo mismo, $\langle\vec{u}, S\vec{v} - T^*\vec{v}\rangle = 0$ para toda $\vec{v}$ en V lo que implica que $S\vec{v} = T^*\vec{v}$ para cualquier $\vec{v}$ en V. □

DEFINICIÓN 2.2.1. El operador T^* definido en el lema 2.2 es llamado el operador adjunto de T con respecto al producto interior $\langle\cdot,\cdot\rangle$.

OBSERVACIÓN 4. En nuestra demostración de la existencia del adjunto, hemos explotado el hecho de que nuestro espacio vectorial tiene una base ortonormal finita. En espacios de dimensión infita este resultado no es válido en toda su generalidad y, como veremos en capítulos posteriores, deberan pedirse condiciones adicionales a un operdor, para que su adjunto este bien definido.

Mostraremos ahora que, la matriz asociada al operador adjunto T^* en una base ortonormal, es la transpuesta de la matriz asociada al operador T.

TEOREMA 2.2.2. *Sea V un espacio vectorial y $T\colon V \to V$ un operador lineal cuya matriz asociada en una base ortonormal $\{\vec{u}_1, \ldots, \vec{u}_n\}$ de V es la matriz $A = (a_{ij})$. Entonces, la matriz asociada al operador adjunto T^*, en esta base, es la matriz $A^t = (b_{ij})$ con $b_{ij} = a_{ji}$ $(i, j = 1, \ldots, n)$.*

DEMOSTRACIÓN. Puesto que $\{\vec{u}_1, \ldots, \vec{u}_n\}$ es una base ortonormal se tiene que $T\vec{u}_j = \sum_{i=1}^{n} \langle T\vec{u}_j, \vec{u}_i\rangle\, \vec{u}_i$ por ende $a_{ij} = \langle T\vec{u}_j, \vec{u}_i\rangle$. De modo análogo se tiene que $b_{ij} = \langle T^*\vec{u}_j, \vec{u}_i\rangle = \langle\vec{u}_i, T^*\vec{u}_j\rangle$. Así, de acuerdo con la definición de adjunto, $b_{ij} = \langle T\vec{u}_i, \vec{u}_j\rangle = a_{ji}$. □

Se puede generalizar la noción de adjunto a operadores cuyo dominio y contradominio no sean el mismo espacio. El procedimiento sería el siguiente: Dado $T : U \to V$, damos bases ortonormales $\mathcal{U}$ y $\mathcal{V}$ en U y V, respectivamente y encontramos la matriz $m \times n$, $A = (a_{ij})$, asociada a T en dichas bases. Definimos A^* como la matriz $n \times m$ cuyos elementos b_{ij} están dados por la relación, $b_{ij} = a_{ji}$. Esta matriz tiene asociado, en las bases $\mathcal{U}$ y $\mathcal{V}$, un operador T^* con dominio V y contradominio U. *El operador T^* es el que, en este caso , definimos como el adjunto de T.* No es difícil verificar que esta definición no depende de la base ortonormal que usemos para representer a T como matriz, pero se requiere de algunas propiedades de las matrices ortogonales que discutiremos más adelante.

De manera alternativa uno puede definir el adijunto de un operador lineal $T : U \to V$, utilizando el hecho de que se puede mostrar que si $\langle\cdot,\cdot\rangle_U$ es el producto inteior en U y $\langle\cdot,\cdot\rangle_V$ es el producto interior en V, entonces, existe un único perador $T^* : V \to U$ tal que

$$\langle T\vec{u}, \vec{v}\rangle_V = \langle\vec{u}, T^*\vec{v}\rangle_U \;\; \forall\vec{u} \in V$$

Por supuesto en este caso el operador T^* definirá al adjunto de T. En el caso de epacios de dimensión finita las dos definiciones son equivalentes.

Un caso particularmente importante de operadores son los operadores que son su propio adjunto. En forma más precisa

DEFINICIÓN 2.2.3. Un operador $T\colon V \to V$ es llamado *auto-adjunto* si y sólo si T es su propio adjunto, esto es,si y sólo si $T = T^*$.

En este caso uno obtiene del teorema 2.2.2 el siguiente resultado.

TEOREMA 2.2.4. *Sea $T\colon V \to V$ es un operador lineal auto-adjunto. Entonces, si V la matriz A asociada a T en cualquier base ortonormal es* ***simetrica,*** *esto es $A = A^t(a_{ij} = a_{ji})$.*

En el capítulo siguiente estudiaremos con detalle las propiedades de estos operadores. Baste, por el momento, señalar que si T es auto-adjunto, entonces $\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, T\vec{v}\rangle$ para todo $\vec{u}$ y $\vec{v}$ en V.

A continuación damos una lista de las propiedades más usadas del adjunto de un operador.

TEOREMA 2.2.5. *Sean $T, S\colon V \to V$ operadores lineales, entonces*

i) *T es no singular si y sólo si T^* es no singular y, en tal caso,* $(T^*)^{-1} = (T^{-1})^*$

ii) $(TS)^* = S^*T^*$

iii) $\det T^* = \det T$

iv) $(\lambda T)^* = \lambda T^*$ *y* $T^* + S^* = (T + S)^*$

2.3. Los operadores de Proyección

En es estudio de espacios vectoriales con producto interior, existen dos clases de operadores particularmente útiles, estos son los operadores de proyección y los operadores ortogonales. En esta sección estudiaremos los primeros.

En lo que resta de este capítulo asumiremos que V es un espacio de dimensión finita con su producto interior $\langle \cdot, \cdot \rangle$.

Sea E un subespacio de V, en el capítulo anterior se estableció que todo $\vec{v} \in V$ tiene una única proyección ortogonal $P_E\vec{v}$ sobre E, luego la relación que a cada $\vec{v}$ le asigna el elemento en V más cercano a $\vec{v}$, define una función de V en V. Dicha función es llamado el operador de proyección o proyección ortogonal de $\vec{v}$ sobre E y la denotamos también por P_E.

Mostraremos ahora que P_E define un operador lineal. En efecto, dados $\vec{u}$ y $\vec{v}$ en V, sabemos que $\vec{u} - P_E\vec{u} \perp E$ y que $\vec{v} - P_E\vec{v} \perp E$ luego

$$\begin{aligned}\langle \alpha\vec{u} + \beta\vec{v} - (\alpha P_E\vec{u} + \beta P_E\vec{v}), \vec{w}\rangle &= \alpha \langle \vec{u} - P_E\vec{u}, \vec{w}\rangle + \beta \langle \vec{v} - P_E\vec{v}, \vec{w}\rangle \\ &= 0 \ \forall \vec{w} \in E\end{aligned}$$

lo que implica que $\alpha\vec{u} + \beta\vec{v} - (\alpha P_E\vec{u} + \beta P_E\vec{v}) \perp E$. Este hecho junto con el teorema de Pitégoras garantizan que para cualquier $\vec{w}$ en E

$$\begin{aligned}\|\alpha\vec{u} + \beta\vec{v} - \vec{w}\|^2 &= \|[\alpha\vec{u} + \beta\vec{v} - (\alpha P_E\vec{u} + \beta P_E\vec{v})] + [(\alpha P_E\vec{u} + \beta P_E\vec{v}) - \vec{w}]\|^2 \\ &= \|\alpha\vec{u} + \beta\vec{v} - (\alpha P_E\vec{u} + \beta P_E\vec{v}\|^2 + \|\alpha P_E\vec{u} + \beta P_E\vec{v}) - \vec{w}\|^2\end{aligned}$$

por tanto

$$\|\alpha\vec{u} + \beta\vec{v} - \vec{w}\| \geq \|\alpha\vec{u} + \beta\vec{v} - (\alpha P_E\vec{u} + \beta P_E\vec{v}\|$$

para todo $\vec{w}$ en E, lo que muestra que $(\alpha P_E\vec{u} + \beta P_E\vec{v})$ minimiza la distancia de $\alpha\vec{u} + \beta\vec{v}$ al subespacio E. Finalmente, de la unicidad de la proyección ortogonal podemos concluir que

$$(\alpha P_E\vec{u} + \beta P_E\vec{v}) = P(\alpha\vec{u} + \beta\vec{v})$$

esto demuestra el teorema

TEOREMA 2.3.1. *Sea P_E la función que a cada $\vec{v}$ en V le asigna su proyección ortogonal sobre E. Entonces, P_E define un operedor lineal de V en V.*

DEFINICIÓN 2.3.2. El operador P_E es llamado el *operador de proyección* u operador proyección sobre E.

Antes de discutir algunas propiedades de los operadores de proyección, introducimos la noción de conjunto ortogonal.

DEFINICIÓN 2.3.3. Dado un conjunto A de V definimos el *conjunto ortogonal* a A como el conjunto de todos los vectores perpendiculares a A y lo denotamos por $A^{\perp}$, esto es,

$$A^{\perp} = \{\vec{v} \in V \mid \langle \vec{v}, \vec{u} \rangle = 0 \ \forall \vec{u} \in A\}$$

$A^{\perp}$ es un subespacio, puesto que si $\vec{v}_1$ y $\vec{v}_2$ están en $A^{\perp}$, entonces $\langle \alpha\vec{v}_1 + \beta\vec{v}_2, \vec{u} \rangle = \alpha \langle \vec{v}_1, \vec{u} \rangle + \beta \langle \vec{v}_2, \vec{u} \rangle = 0$ para toda $\vec{u}$ en A. Obsérvese que todo $\vec{u}$ en A es perpendicular a $A^{\perp}$.

El siguiente teorema es, en gran medida, consecuencia de las propiedades geométricas de las proyecciones ortogonales discutidas en el capitulo 1.

TEOREMA 2.3.4. *Sea E un espacio de V y sea P_E el operador de proyección sobre E, entonces,*

i) $P_E\vec{u} = \vec{u} \iff \vec{u} \in E$. *Por lo tanto, el rango de P_E es E.*
ii) $\ker P_E = E^{\perp}$.
iii) $I - P_E$ *es el operador de proyección sobre* $E^{\perp}$.
iv) P_E *es idempotente, i.e. ,*$P_E^2 = P_E$.
v) $P_E^* = P_E$, *esto es* $\langle P_E\vec{u}, \vec{v} \rangle = \langle \vec{u}, P_E\vec{v} \rangle \ \forall \vec{u}, \vec{v} \in V$.

Demostración. El inciso i) se sigue del hecho de que $\vec{u} - P_E\vec{u} \in E$ y $\vec{u} - P_E\vec{u} \perp E$ y se deja como ejercicio. Como consecuencia de este reultado se obtiene iv). Quedan por demostrar ii), iii) y v).

ii) Sabemos que $\vec{v} - P_E\vec{v} \perp E$. Si $\vec{v} \in \ker P_E$, se tiene que $\vec{v} - P_E\vec{v} = \vec{v}$ y por ende $\vec{v} \perp E$. Recíprocamente, si $\vec{v}$ esta en $E^\perp$,

$$0 = \langle \vec{v} - P_E\vec{v}, \vec{u} \rangle = \langle \vec{v}, \vec{u} \rangle - \langle P_E\vec{v}, \vec{u} \rangle = -\langle P_E\vec{v}, \vec{u} \rangle \ \ \forall \vec{u} \in E$$

en particular para $\vec{u} = P_E\vec{v}$, luego $-\|P_E\vec{v}\|^2 = -\langle P_E\vec{v}, P_E\vec{v} \rangle = 0$ y por lo tanto $\vec{v} \in \ker P_E\vec{v}$.

iii) Puesto que $(I - P_E)\vec{w} = \vec{w} - P_E\vec{w}$ y $\vec{w} - P_E\vec{w} \perp E$, se tiene que $\vec{w} - P_E\vec{w} \in E^\perp$. Por otra parte $\vec{w} - (I - P_E)\vec{w} = P_E\vec{w} \perp E^\perp$, así podemos aplicar el teorema de Pitagoraspara concluir que

$$(I - P_E)\vec{w} = P_{E^\perp}\left(\vec{w}\right) \ \ \forall \vec{w} \in V$$

Falta demostrar v), para ello, empecemos por observar que para toda $\vec{u}$ y $\vec{v} \in V$

$$\vec{u} = P_E\vec{u} + (I - P_E)\vec{u} \ \ \text{y} \ \ \vec{v} = P_E\vec{v} + (I - P_E)\vec{v}$$

De esto, obtenemos las identidades

$$\begin{aligned}
\langle P_E\vec{u}, \vec{v} \rangle &= \langle P_E\vec{u}, P_E\vec{v} + (I - P_E)\vec{v} \rangle = \langle P_E\vec{u}, P_E\vec{v} \rangle + \langle P_E\vec{u}, (I - P_E)\vec{v} \rangle \\
\langle \vec{u}, P_E\vec{v} \rangle &= \langle P_E\vec{u} + (I - P_E)\vec{u}, P_E\vec{v} \rangle = \langle P_E\vec{u}, P_E\vec{v} \rangle + \langle (I - P_E)\vec{u}, P_E\vec{v} \rangle
\end{aligned}$$

Finalmente,como el $\Im P_E = E$ y $\Im(I - P_E) = E^\perp$,

$$\langle P_E\vec{u}, (I - P_E)\vec{v} \rangle = \langle (I - P_E)\vec{u}, P_E\vec{v} \rangle = 0$$

por tanto,

$$\langle P_E\vec{u}, \vec{v} \rangle = \langle P_E\vec{u}, P_E\vec{v} \rangle = \langle \vec{u}, P_E\vec{v} \rangle \quad \forall \vec{u}, \vec{v} \in V \qquad \square$$

Las propiedades iv) y v) enunciadas en el teorema anterior, caracterizan completamente a los operadores de proyección en espacios de dimensión finita en el siguiente sentido:

Teorema 2.3.5. *Si $P\colon V \to V$ es un operador lineal tal que $P^2 = P$ y $P^* = P$, entoces, P es el operador de proyección sobre el subespacio $E = \Im P$.*

Demostración. Sea $E = \Im P$, mostraremos que $(\vec{v} - P\vec{v}) \perp E$ para todo $\vec{v} \in V$. Sabemosque para todo $\vec{u}$ en E existe al menos un $\vec{w}$ en V tal que $P\vec{w} = \vec{u}$, así que

$$\begin{aligned}
\langle \vec{v} - P\vec{v}, \vec{u} \rangle = \langle \vec{v} - P\vec{v}, P\vec{w} \rangle &= \langle P^*(\vec{v} - P\vec{v}), \vec{w} \rangle \\
&= \langle P(\vec{v} - P\vec{v}), \vec{w} \rangle = \left\langle P\vec{v} - P^2\vec{v}, \vec{w} \right\rangle \\
&= \langle P\vec{v} - P\vec{v}, \vec{w} \rangle = 0
\end{aligned}$$

Para la tercera igualdad hemos usado el hecho de que $P = P^*$, mientras que para la penúltima igualdad se aplicó la idempotencia de P.

Como consecuencia, tenemos que $(\vec{v} - P\vec{v}) \perp (P\vec{v} - \vec{u})$. Por el teorema de Pitágoras

$$\begin{aligned}\|\vec{v} - \vec{u}\|^2 &= \|(\vec{v} - P\vec{v}) + (P\vec{v} - \vec{u})\|^2 \\ &= \|(\vec{v} - P\vec{v})\|^2 + \|(P\vec{v} - \vec{u})\|^2 \geq \|\vec{v} - P\vec{v}\|^2\end{aligned}$$

para todo $\vec{u}$ en E. Lo que muestra que $P\vec{v}$ minimiza la distancia de $\vec{v}$ a $E = \Im P$ y por ende, P es el operador de proyección sobre E. $\square$

Como consecuencia de estos resultados sobre proyecciones, uno tiene de manera inmediata que $E \cap E^{\perp} = \{0\}$ y que todo elemnto de $\vec{v} \in V$ se puede escribir de manera única como elementos de E más un elemento de $E^{\perp}$. La primera afirmación se sigue del hecho de que $E = \ker P$ y $E^{\perp} = \ker(I - P)$. Por tanto, si $\vec{u} \in E \cap E^{\perp}$, $P\vec{u} = 0$ y $\vec{u} - P\vec{u} = 0$, lo que implica que $\vec{u} = 0$. En cuanto a la segunda afirmación es claro que $\vec{v} = P\vec{v} + (I - P)\vec{v}$ para todo $\vec{v} \in V$, así, todo elmento de V se puede expresar como suma de un elemento de E más un elemento de $E^{\perp}$. Para ver que dicha expresión es única basta notar que si $\vec{u} + \vec{u}^{\perp} = P\vec{v} + (I - P)\vec{v}$ con $\vec{u} \in E$ y $\vec{u}^{\perp} \in E^{\perp}$, entonces,

$$\vec{u} - P\vec{v} = (I - P)\vec{v} - \vec{u}^{\perp}$$

Pero $\vec{u} - P\vec{v} \in E$ y $(I - P)\vec{v} - \vec{u}^{\perp} \in E^{\perp}$, y $E \cap E^{\perp} = \{0\}$, por lo que, necesariamente

$$\vec{u} - P\vec{v} = (I - P)\vec{v} - \vec{u}^{\perp} = 0$$

TEOREMA 2.3.6. *Sea V un espacio vectorial y sea E un subespacio de V. Entonces, $E \cap E^{\perp} = \{0\}$ y todo elemnto de V se puede expresar de manera única, como la suma de un elemento de E más un elemento de $E^{\perp}$. Dichos elementos son las proyecciones sobre E y $E^{\perp}$.*

Este resultado nos permite descomponer a nuestro espacio V en dos subespacios de dimensión menor y por ende menos difíciles de trabajar. Nosotros ahora introducimos el consepto de suma directa de subespacios

DEFINICIÓN 2.3.7. Dado un espacio vectorial V y subespacios E y F diremos que V es *la suma directa* de E y F si se cumplen las dos condiciones siguientes:

i) $E \cap F = \{0\}$

ii) Todo elemeneto de V se escriba como suma de un elemento de E más un elemento de E más un elemento de F. En tal caso escribimos $V = E \oplus F$.

En forma más general, si $E_1, \ldots, E_k$ son subespacios de V tales que $E_j \cap E_i = \{0\}$ y todo elemento $\vec{v}$ de V se puede escribir de la forma

$$\vec{v} = \vec{u}_1 + \cdots + \vec{u}_k \text{ con } \vec{u}_i \in E_i$$

diremos que V se descompone como suma directa de $E_1, \ldots, E_k$ y escribiremos $V = E_1 \oplus \cdots \oplus E_k$.

Daremos ahora una lista de propiedades de los operadores de proyección cuya demostración se deja como ejercicio.

TEOREMA 2.3.8. *Sea $(V, \langle\cdot\,\cdot\rangle)$ un espacio vectorial con producto interior y sean P y Q operadores de proyecciones, entonces*

i) $\langle P\vec{u}, \vec{u}\rangle = \|P\vec{u}\|^2, \forall \vec{u} \in V$
ii) $\|P\vec{u}\| = \|\vec{u}\|$ *si y sólo si* $\vec{u} \in \Im P$.
iii) $\Im P \perp \Im Q$ *si y sólo si* $PQ \equiv 0$, *en este caso escribiremos* $P \perp Q$
iv) $P + Q$ *es un operador de proyección si y sólo si* $PQ \equiv 0$
v) PQ *es un operador de proyección si y sólo si* $PQ \equiv QP$
vi) $P - Q$ *es un operador de proyección si y sólo si* $\Im Q \subset \Im P$

Un resultado que nos sera de utilidad más adelante y que se desprende de este teorema es el siguiente:

TEOREMA 2.3.9. *Sean $P_1, \ldots, P_k$ proyecciones sobre V y sea $P = \sum_{i=1}^{n} P_i$, entonces, P es una proyección si y sólo si $P_i \perp P_j$ para $i \neq j$ .Además, si P es proyección, entonces,*

$$\Im P = \Im P_1 \oplus \cdots \oplus \Im P_k$$

DEMOSTRACIÓN. Si la familia de $\{P_i\}$ es ortogonal, *i.e.*, $P_i \perp P_j$ para cada $i \neq j$,se tiene, del inciso iii) del teorema anterior que

$$P^2 = (\sum_{i=1}^{k} P_i)(\sum_{i=1}^{k} P_i) = \sum_{i=1}^{k} P_i^2 = \sum_{i=1}^{k} P_i = P$$

Por otra parte

$$\langle P\vec{u}, \vec{v}\rangle = \left\langle \sum_{i=1}^{k} P_i\vec{u}, \vec{v} \right\rangle = \left\langle \vec{u}, \sum_{i=1}^{k} P_i\vec{v} \right\rangle = \langle \vec{u}, P\vec{v}\rangle$$

para toda $\vec{u}$ y $\vec{v}$ en V lo que muestra que $P = P^*$.

Si P es proyección y $\vec{u}$ esta en el $\Im P_j$, se tiene que

$$\|\vec{u}\|^2 \geq \|P\vec{u}\|^2 = \langle P\vec{u}, \vec{u}\rangle = \sum_{i=1}^{k} \langle P_i\vec{u}, \vec{u}\rangle = \sum_{i=1}^{k} \|P_i\vec{u}\|^2 \geq \|P_j\vec{u}\|^2 = \|\vec{u}\|^2$$

Esto implica, por una parte, que $P_i\vec{u} = 0$, $i \neq j$, y en consecuencia $\vec{u} \perp \Im P_i$ (teorema 2.3.1-ii)) y por la otra, que $\vec{u}$ esta en el rango de P. Así, $P_i \perp P_j$ para $i \neq j$ y $\Im P_i \subset \Im P$. Mas aún, puesto que los rangos de las P_i son mutuamente perpendiculares, su suma diercta está bien definida y

$$\Im P_1 \oplus ..... \oplus \Im P_k \subset\subset \Im P$$

Pero si $\vec{u} \in \Im P$, entonces,

$$\vec{u} = P\vec{u} = \sum_{i=1}^{k} P_i\vec{u} \in \Im P_1 \oplus \cdots \oplus \Im P_k$$

lo que muestra que el rango de P es la uma directa de los rangos de las P_i. □

No está por demás insistir en que, para definir el operador de proyección, hemos usado implícitamente la hipótesis de que V es un espacio de dimensión finita. Para espacios más generales, habrá que imponer condiciones adicionales al subespacio E y al operador P si queremos que los teoremas 2.3.6 a 2.3.9 sean válidos.

2.4. Operadores Ortogonales y unitarios

La clase de operadores que preservan la métrica euclidiana, no sólo son de interés en geometría, su utilidad en análisis, álgebra y, las aplicaciones en general, lo hacen particulamente importantes. En esta sección daremos una caracterización de estos operadores en espacios vectoriales de dimensión finita y mostraremos algunas de las propiedades de las matrices asociadas a dichos operadores.

Empecemos por recordar que si V es un espacio con producto interior, entonces, V tiene una norma asociada con dicho produto interior. La norma esta dada por la fórmula $\|\vec{u}\|^2 = \langle \vec{u}, \vec{u} \rangle$.

TEOREMA 2.4.1. *Sea $R\colon V \to V$ un operador lineal, entonces, las siguientes proposiciones son equivalentes:*

i) $\|R\vec{u}\| = \|\vec{u}\|$ *paratoda $\vec{u}$ en V*
ii) $\langle R\vec{u}, R\vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle$ *para toda $\vec{u}$ y $\vec{v}$ en V*
iii) $R^*R = I$ *y* $RR^* = I$
iv) *R es invertible y* $R^{-1} = R^*$

DEMOSTRACIÓN. $i) \Rightarrow ii)$ Pusto que $4\langle \vec{u}, \vec{v} \rangle = \|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2$ y

$$\|R\vec{u} + R\vec{v}\| = \|\vec{u} + \vec{v}\| \text{ y } \|R\vec{u} - R\vec{v}\| = \|\vec{u} - \vec{v}\|$$

se tiene que

$$4\langle \vec{u}, \vec{v} \rangle = \|R\vec{u} + R\vec{v}\|^2 - \|R\vec{u} - R\vec{v}\|^2$$

Si desarrolamos las normas que aparecen en la derecha de la igualdad, enterminos del producto interior obtenemos que

$$\|R\vec{u} + R\vec{v}\|^2 - \|R\vec{u} - R\vec{v}\|^2 = 4\langle R\vec{u}, R\vec{v} \rangle$$

y por ende

$$4\langle \vec{u}, \vec{v} \rangle = 4\langle R\vec{u}, R\vec{v} \rangle$$

$ii) \Rightarrow iii)$ Deacuerdo con la definición de adjunto, se tiene de ii) que $\langle R^*R\vec{u}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle$ para toda $\vec{u}$, $\vec{v} \in V$ en consecuencia para cada $\vec{u} \in V$,

$\langle R^*R\vec{u} - \vec{u}, \vec{v}\rangle = 0$ para toda $\vec{v} \in V$ lo que implica que $R^*R\vec{u} - \vec{u} = 0$ para todo $\vec{u} \in V$, por lo tanto, $R^*R = I$. De modo similar uno obtiene que $RR^* = I$.

$iii) \Rightarrow iv)$ Se sigue de la deficición de operador inverso.

$iv) \Rightarrow i)$ Puesto que R es invertible, $\left\langle R^{-1}R\vec{u}, \vec{u}\right\rangle = \langle \vec{u}, \vec{u}\rangle = \|\vec{u}\|^2$. Por otra parte, como $R^{-1} = R^*$

$$\left\langle R^{-1}R\vec{u}, \vec{u}\right\rangle = \langle R^*R\vec{u}, \vec{u}\rangle = \langle R\vec{u}, R\vec{u}\rangle = \|R\vec{u}\|^2$$

Lo que muestra que $\|R\vec{u}\|^2 = \|\vec{u}\|$. □

DEFINICIÓN 2.4.2. Un operador $R : V \to V$ para el cual se cumple la condición

$$\|R\vec{u}\| = \|\vec{u}\| \ \ \forall u \in V$$

es llamado *un operador ortogonal* en V.

El hecho de que $R^*R = I$, da una caracteristica particular a la matriz asociada a un operador ortogonal o unitario $R : V \to V$ en cualquier base ortogonal. A saber las columnas de la matriz son mutuamemte parependiculares y de longitud 1. En efecto, si denotamos por O la matriz asociada a un operador ortogonal y si $\vec{w}_1, \ldots, \vec{w}_n$ son los vectores columna de O, entonces, el elemento ι_{ij} de la matriz O^*O está dado por $\iota_{ij} = \sum_{k=1}^{n} w_{ik}^* w_{jk}$, donde w_{ik} es la componente k del vector columna $\vec{w}_i$ y w_{jk} es el elemento k del vector columna $\vec{w}_i$ y w_{ik} es el elemento de k del ve ctor columna $\vec{w}_j$. Así, si O es la matriz asociada R en una base ortorormal, se tiene que $O^*O = I$ y por ende $\iota_{ij} = \vec{w}_i \cdot \vec{w}_j = \delta_{ij}$, donde

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Lo que muestra que las columnas de O forman una base ortonormal para $\mathbb{R}^n$. Una matriz con esta propiedad es llamada una *matriz ortogonal.*

TEOREMA 2.4.3. *Un operador es ortogonal si y sólo si la matriz asociada a ese operador en alguna base ortonormal es una matriz ortogona, esto es, si y sólo si las columnas de la matriz asociada forman una base ortonormal.*

En el caso del espacio euclidiano una matriz ortogonal O puede interpretarse como una rotación de los ejes coordenados, la cual transforma dichos ejes en los ejes determinados por los vectores columna de la matriz O y por tanto como un cambio de cooordenadas rectangulares. De la misma manrera, en el caso de un espacio con producto interior, uno puede pensar a un operador ortogonal como un cambio de coordenadas que preserva la norma y el producto interior.

2.5. Problemas

1. Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio de dimensión finita con producto interior.

a) A un operador lineal definido en V, cuyo contradominio er $\mathbb{R}$ se le llama *función lineal.* Muestre que el conjunto funciones lineales de V es un espacio vectorial. Este espacio vectorial es llamado el *dual de* V y se denota por V^*.

b) Muestre que, toda funcional lineal l de V puede escribirse en la forma $l(\vec{u}) = \langle\bar{a}, \vec{u}\rangle$ con $\bar{a}$ un vector en V y, recíprocamente, toda función de la forma $l(\vec{u}) = \langle\bar{a}, \vec{u}\rangle$ es una función lineal de V. (Sug.: Use la representación matricial de f)

c) Mustre que el vector $\bar{a}$ asociado a l en el inciso b) es único y que dicha asociación define un isomorfismo de espacios vectoriales.

d) Muestre que el $\ker l$ es el subespacio de todos los vectores ortogonales al vector $\bar{a}$.

2. Considere el espacio $P_n[\cos x]$ del ejemplo 1.3 del Capítulo 1.

a) Muestre que el operador $Tu = \frac{d^2u}{dx^2} - u$ es un operador lineal.

b) Para $n = 3$, encuentre la matriz asociada a T en la base$\{1, \cos x, \cos^2 x, \cos^3 x\}$.

c) Para $n = 3$, encuentre la matriz asociada a T en la base$\{1, \cos x, \cos 2x, \cos 3x\}$

d) Diga si T tiene inversa. En caso afirmativo diga cual es.

3. Considere el espacio $P_n[x]$ del ejemplo 1.4 del Capítulo I

a) Muestre que el operador

$$Lu = (x^2 - 1)\frac{d^2u}{dx^2} + 2x\frac{du}{dx}$$

es un operador lineal.

b) Para $n = 3$, encuentre la matriz asociada a L en la base $\{1, x, x^2, x^3\}$.

c) Para $n = 3$ encuentre la matriz asociada a L en la base dada por los polinomios de Legendre.

d) Diga si L tiene inversa. En caso afirmativo calcule cual es.

4. Muestre que si $\mathcal{U}$ y $\mathcal{V}$ son bases ortonormales de U y V, respectivamente, y $T : U \to V$ es lineal, entonces, la matriz asociada a T en estas bases tinen como elementos a_{ij} al valor $\langle T\vec{u}_j, \vec{v}_i\rangle$.

5. Encuentre el adjunto del operador T del problema 2 directamente de la definición. (Sug.: Integre por partes y observe que todo elemento de $P_n[\cos x]$ se anulan en $\pm\frac{\pi}{2}$).

6. Encuentre el adjunto del operador L del problema 3 directamente de la definicinición de adjunto.

7. Sea$\big(V, \langle\cdot,\cdot\rangle\big)$ un espacio vectorial con producto interior. Muestre que, módulo el isomorfismo definido en el inciso c) del problema 1, una definición alternativa del adjunto de un operador es la siguiente:

El operador adjunto de T *es el operador* $\widetilde{T}^*$ *que a cada funcional* l *le asigna la funcional lineal* $\widetilde{T}^* l$ *dada por la relación* $\left(\widetilde{T}^* l\right)(\vec{u}) = l(T\vec{u})$. Esto es, muestre que si $\iota \colon V^* \to V$ es el isomorfismo definido en el inciso c) del problema 1, entonces, $T^* = \widetilde{T}^* \circ \iota$

8. Sea $T\colon U \to V$ $(U \neq V)$ un operador lineal, muestre que $T^*\colon V \to U$ es el adjunto de T si y sólo si

$$\langle T\vec{u}, \vec{v}\rangle_V = \langle \vec{u}, T^*\vec{v}\rangle_U \ \ \forall \vec{u} \in U \text{ y } \forall \vec{v} \in V$$

9. Demuestre el teorema 2.2.4

10. Termine la demostración del teorema 2.3.1-*iii*)

11. Demuestre el teorema 2.3.6

Capítulo 3

Teorema Espectral en Espacios de Dimensión Finita

3.1. Formas cuadráticas y Familias de Cónicas

En el estuidio de las cónicas aparecen, de modo natural, expresiones de la forma

$$Q[x_1, x_2] = ax_1^2 + abx_1x_2 + cx_2^2$$

las cuales son llamadas formas cuadráticas en dos variables. En efecto, uno estudia en sus cursos de geometría que cualquier hipérbola o elipse con centro en el origen satisface una ecuación de la forma

$$ax_1^2 + abx_1x_2 + cx_2^2 = k \tag{3.1}$$

y si aceptamos que el vacío, un punto o un par de rectas son casos degenerados de cónicas, entonces, para cualquier valor de k la ecuación 3.1 representa una cónica. Así, uno puede pensar que una forma cuadrática en dos variables represente una familia de cónicas.

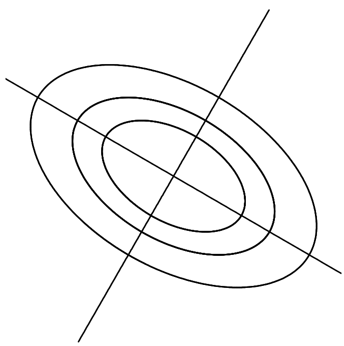

Figure 3.1

Es bien sabido que mediante una rotación de los ejes coordenados a los ejes principales de la cónica, una ecuación del tipo 3.1 puede reducirce a la forma canónica

$$\lambda_1 {x'_1}^2 + \lambda_2 {x'_2}^2 = k$$

En particular, uno tiene que para cualquier valor de k la misma rotación reduce la ecuación 3.1 a la forma canónica y, por lo tanto, todas las cónicas de la familia determinada por 3.1, tienen los mismos ejes principales.

En el caso de superficies cuadráticas la situación es muy similar; para cada valor de k una ecuación de la forma

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 = k \qquad (3.2)$$

representa un elipsoide, un hiperboloide de una o dos hojas, un cono o a cualquiera de los posibles casos degenerados. Más aún, para diferentes valores de k la superficies cuadráticas que se obtienen tienen los mismos ejes principales.

También en este caso, una rotación de los ejes coordenados a los ejes principales de la superficies determinada por 3.2 reduce nuestra ecuación a su forma canónica

$$\lambda_1 {x'_1}^2 + \lambda_2 {x'_2}^2 + \lambda {x'_3}^2 = k$$

Así, una función de la forma

$$Q[x_1, x_2, x_3] = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$

la cual es llamada *una forma caudrática en tres variables*, puede interpretarse geométricamente como una familia de superficies cuadráticas, todas ellas con los mismos ejes principales.

3.2. Reducción de Formas Cuadráticas a su Forma Canónica. Teorema Espectral para formas Cuadráticas

Nosotros generalizaremos ahora las ideas geométricas discutidas en la sección precedente. Empezaremos por dar una definición precisa de lo que es una forma cuadrática en n variables.

DEFINICIÓN 3.2.1. Una fnción $Q\left[\vec{x}\right] = Q\left[x_1, \ldots, x_n\right]$ de la forma

$$Q\left[\vec{x}\right] = Q\left[x_1, \ldots, x_n\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

es llamada una **forma cuadrática** en E^n.

Dos observaciones inmediatas de esta definición son:

PROPOSICIÓN 3.2.2.

i) *Una forma cuadrática es homogénea de grado 2, esto es*

$$Q\left[\mu\vec{x}\right] = \mu^2 Q\left[\vec{x}\right], \ \forall \mu \in \mathbb{R}$$

ii) Siempre podemos escribir una forma cuadrática de manera simétrica,esto es, de modo que $a_{ij} = a_{ji}$. (Basta tomar $\frac{1}{2}(a_{ij} + a_{ji})$ en lugar de a_{ij} y de a_{ji}.)

Escribir una forma cuadrática de manera simétrica, permite caracterizar a dicha forma en terminos de sus coeficientes.

NOTACIÓN 4. En lo que sigue resultará conveniente denotar por $\vec{e}_1, \dots, \vec{e}_n$ a los vectores canónicas en E^n, esto es, $\vec{e}_1 = [1, 0, \dots, 0], \vec{e}_2 = [0, 1, \dots, 0], \vec{e}_3 = [0, 0, 1, \dots, 0]$, etc.

PROPOSICIÓN 3.2.3. *Suponga que las formas cuadráticas Q_1 y Q_2 se escriben en la forma*

$$Q_1\left[\vec{x}\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad \text{con } a_{ij} = a_{ji}$$

y

$$Q_2\left[\vec{x}\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j \text{ con } b_{ij} = b_{ji}$$

Entonces, $Q_1 \equiv Q_2$ si y sólo si $a_{ij} = b_{ij}$ para toda i y j entre 1 y n.

En otra términos, toda forma cuadrática está caracterizada de manera única por sus coefeicientes en la representación simétrica

DEMOSTRACIÓN. Es claro que si $a_{ij} = b_{ij}$, $Q_1 \equiv Q_2$. Para ver el recíproco notemos que

$$Q_1\left[\vec{e}_i + \vec{e}_j\right] = Q_1[0, \dots, \overbrace{1}^{i}, \dots, \overbrace{1}^{j}, \dots, 0] = a_{ii} + 2a_{ij} + a_{jj}$$

$$Q_2\left[\vec{e}_i + \vec{e}_j\right] = Q_2[0, \dots, \overbrace{1}^{i}, \dots, \overbrace{1}^{j}, \dots, 0] = b_{ii} + 2b_{ij} + b_{jj}$$

y como $Q_1[\vec{e}_k] = a_{kk}$ y $Q_2\left[\vec{e}_k\right] = b_{kk}$ se tiene que si $Q_1 \equiv Q_2$ entonces, $a_{ij} = b_{ij}$. □

En la sección precedente hemos establecido que una forma cuadrática en dos o tres variables represente una familia de curvas o superficies cuadráticas con los mismos ejes principales y que dichas funciones pueden reducirse a la forma $\lambda_1 {x'_1}^2 + \lambda_2 {x'_2}^2$ o a la forma $\lambda_1 {x'_1}^2 + \lambda_2 {x'_2}^2 + \lambda_3 {x'_3}^2$ mediante una rotación. Ahora bien, dada una forma cuadrática Q en n variables se puede pensar que Q representa una familia de "hipersuperficies cuadráticas" cuyos "ejes principales" son mutuamente perpendiculares y que, mediante una rotación de los ejes coordenados a los ejes de la familia, se puede reducir la forma cuadrática a una expresión del tipo

$$\lambda_1 {x'_1}^2 + \cdots + \lambda_n {x'_n}^2 \tag{3.3}$$

Sin embargo, no parece tarea fácil (y no lo es) encontrar los supuestos ejes principales y determinar así la rotación que reduce Q a su forma canónica. Afortunadamente, uno puede reducir el problema a uno de máximos y mínimos con restricciones que, módulo el cálculo de varias variables, es fácilmente generalizable. La base para esta reducción son las siguientes observaciones para los casos de dos o tres variables:

OBSERVACIÓN 5. Los ejes principales de la familia de cónicas son mutuamente perpendiculares.

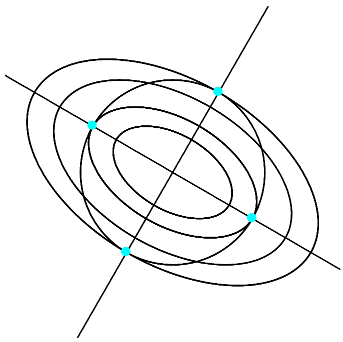

Figure 3.2

OBSERVACIÓN 6. Las direcciones de los ejes principales de la familia de cónicas pueden obtenerse en terminos del problema de máximos y mínimos para lo forma cuadrática restringida al circulo o esfera unitaria. En efecto, uno puede ver (figura 3.2) gráficamente que los vectores de posición de los puntos donde la familia de cónicas es tangente al círculo o esfera de radio 1 están sobre las direcciones principales y dichos puntos son justamente los puntos críticos del problema con restricciones:

$$\begin{cases} Q\left[\vec{x}\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \\ \|\vec{x}\|^2 = \sum_{i=1}^{n} x_i^2 = 1 \end{cases}$$

Partiendo de estas observaciones nosotros debemos elegir como primer eje principal de la familia de hipersuperficies asociada a una forma cuadrática Q en n variables, al vector de posición $\vec{u}_1$ del punto

$(u_{11}, ..., u_{n1})$ tal que $\|\vec{u}_1\|^2 = 1$ y

$$Q\left[\vec{u}_1\right] = Q\left[u_{11}, ..., u_{n1}\right] = Max\left\{Q\left[\vec{x}\right] : \|\vec{x}\|^2 = 1\right\} \tag{3.4}$$

El segundo eje debe ser perpendicular a $\vec{u}_1$ y tambien debe ser un punto crítico, así que el candidato sería el vector $\vec{u}_2 = (u_{12}, ..., u_{n2})$ tal que $\|\vec{u}_2\|^2 = 1$, $\vec{u}_2 \perp \vec{u}_1$ y

$$Q\left[\vec{u}_2\right] = Q\left[u_{12}, ..., u_{n2}\right] = Max\left\{Q\left[\vec{x}\right] : \|\vec{x}\|^2 = 1 \text{ y } \vec{x} \perp \vec{u}_1\right\} \tag{3.5}$$

Siguiendo este razonamiento, el j-ésimo eje principal sería el vector $\vec{u}_j = (u_{1j}, ..., u_{nj})$ tal que $\|\vec{u}_j\|^2 = 1$, $\vec{u}_j \perp \left[\left[\vec{u}_1, ..., \vec{u}_{j-1}\right]\right]$ y

$$Q\left[\vec{u}_j\right] = Q\left[u_{1j}, ..., u_{nj}\right] = Max\left\{Q\left[\vec{x}\right] : \|\vec{x}\|^2 = 1 \text{ y } \vec{x} \perp \left[\left[\vec{u}_1, ..., \vec{u}_{j-1}\right]\right]\right\} \tag{3.6}$$

(Recuerde que $\left[\left[\vec{u}_1, ..., \vec{u}_{j-1}\right]\right]$ denota el subespacio generado por $\vec{u}_1, ..., \vec{u}_{j-1}$)

Mostraremos que, en efecto, el cambio de coordenadas $\vec{x} = R\vec{x}'$, donde R es la matriz ortogonal

$$R = \left(\vec{u}_1 \left|\vec{u}_2\right| \cdots \left|\vec{u}_n\right.\right) = \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & & \vdots \\ u_{1n} & \cdots & u_{nn} \end{pmatrix} \tag{3.7}$$

cuyas columnas son los componentes de los vectores $\vec{u}_1, ..., \vec{u}_n$ que acabamos de construir, nos da el cambio de coordenadas que reduce Q a su forma canónica.

TEOREMA 3.2.4. *(Teorema Espectral versión I):Toda forma cuadrática*

$$Q\left[\vec{x}\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \; (a_{ij} = a_{ji})$$

puede reducirse a la forma

$$Q'\left[\vec{x}'\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i x_i'^2$$

mediante un cambio de coordenadas ortogonales. Esto es, existe un matriz ortogonal R tal que

$$Q'\left[\vec{x}'\right] = Q\left[R\vec{x}'\right] = \sum_{i=1}^{n} \lambda_i x_i'^2$$

DEMOSTRACIÓN. Sean $\vec{u}_1, ..., \vec{u}_n$ los vectores obtenidos vía las relaciones (3.4) y (3.6). Sea R la matriz (3.7) y sea Q' la función dad por el cambio de coordenadas $\vec{x} = R\vec{x}'$, esto es

$$Q'\left[x'\right] = Q\left[R\vec{x}'\right]$$

No es dicícil mostrar que Q' es una forma cuadrática en las variables $(x'_1, ..., x'_n)$ y por ende puede escribirse en la forma

$$Q'\left[x'_1, ..., x'_n\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x'_i x'_j \ (b_{ij} = b_{ji})$$

Así, para ver que Q' está en forma canónica basta demostrar que $b_{ij} = 0$ para $i \neq j$.

Sea $\lambda_1 = Max\left\{Q\left[\vec{x}\right] : \|\vec{x}\|^2 = 1\right\}$. Definimos la forma cuadrática Q_1 como

$$Q_1\left[\vec{x}'\right] = Q'\left[\vec{x}'\right] - \lambda_1(x_1'^2, ..., x_n'^2) = Q'\left[\vec{x}'\right] - \lambda_1 \|\vec{x}\|^2$$

Mostraremos primero que $Q_1\left[\vec{x}'\right] \leq 0$ para toda $\vec{x} \in E^n$. En efecto,

$$Q_1\left[\vec{x}'\right] = Q\left[R\vec{x}'\right] - \lambda_1 \|\vec{x}\|^2$$

Así, para todo $\vec{x}'$ tal que $\|\vec{x}'\|^2 = 1$ se tiene que $Q_1\left[\vec{x}'\right] = Q\left[R\vec{x}'\right] - \lambda_1$ y como R es ortogonal, $\|R\vec{x}'\|^2 = \|\vec{x}'\|^2 = 1$ lo que implica, de acuerdo con la definición de λ_1 que $Q_1\left[\vec{x}'\right] \leq 0$ para todo $\vec{x}'$ de norma 1. Ahora bien, si $\vec{x}' \neq 0$ entonces $\vec{x}'/\|\vec{x}'\|$ es de norma 1 y como toda forma cuadratica es homogénea de grado 2

$$Q_1\left[\vec{x}'\right] = \|\vec{x}'\|^2 Q_1\left[\vec{x}'/\|\vec{x}'\|\right] \leq 0$$

Puesto que el caso $\vec{x}' = 0$ es trivial, necesariamente Q_1 es negativa.

A continuación mostraremos que esto obliga a que los coeficientes b_{1j} y b_{j1} son cero para $j \neq 1$. Empecemos por observar que como $\vec{u}_1 = R\vec{e}_1$, se tiene que $Q'\left[\vec{e}_1\right] = Q\left[R\vec{e}_1\right] = Q\left[\vec{u}_1\right] = \lambda_1$ y por ende $b_{11} = \lambda_1$. Por otra parte, para cualquier ε

$$\begin{aligned} Q'_1\left[\vec{e}_1 + \varepsilon\vec{e}_j\right] &= b_{11} + b_{1j}\varepsilon + b_{j1}\varepsilon + b_{jj}\varepsilon^2 - \lambda_1 - \lambda_1\varepsilon^2 \\ &= (b_{1j} + b_{j1})\varepsilon + (b_{jj} - \lambda_1)\varepsilon^2 \\ &= 2b_{1j}\varepsilon + (b_{jj} - \lambda_1)\varepsilon^2 \leq 0 \end{aligned}$$

pero esto sólo puede ocurrir si $b_{1j} = 0$. Como $b_{j1} = b_{1j}$ obtenemos que, en efecto, los b_{1j} y b_{j1}, $j \neq 1$, son todos cero.

De esto concluimos que

$$Q'\left[\vec{x}'\right] = \lambda_1 x_1'^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}x'_i x'_j$$

Tomemos ahora

$$\lambda_2 = Max\left\{Q\left[\vec{x}\right] : \|\vec{x}\|^2 = 1 \text{ y } \vec{x}\perp\left[\left[\vec{u}_1\right]\right]\right\}$$

y definamos Q'_2 como

$$Q'_2\left[\vec{x}'\right] = Q'\left[\vec{x}'\right] - \lambda_2(x_2'^2 + ... + x_n'^2)$$

Mostraremos que $Q_2'\left[\vec{x}'\right] \leq 0$ para todo $\vec{x}\perp\left[\left[\vec{e}_1\right]\right]$, esto es, si $\vec{x}' = (0, x_2', ..., x_n')$. En este caso se tiene que

$$Q_2'\left[\vec{x}'\right] = Q'\left[\vec{x}'\right] - \lambda_2 \left\|\vec{x}'\right\|^2 \text{ para todo } \vec{x}\perp\left[\left[\vec{e}_1\right]\right]$$

Nuevamente, si $\left\|\vec{x}'\right\|^2 = 1$ y $\vec{x}\perp\left[\left[\vec{e}_1\right]\right]$, $Q_2'\left[\vec{x}'\right] = Q\left[R\vec{x}'\right] - \lambda_2$. Ya que R es una matriz ortogonal,$\left\|R\vec{x}'\right\|^2 = 1$ y $R\vec{x}\perp\left[\left[\vec{e}_1\right]\right] = \left[\left[\vec{u}_1\right]\right]$. De la definición de λ_2 concluimos que $Q_2'\left[\vec{x}'\right] \leq 0$ para todo $\vec{x}'$ tal que $\left\|\vec{x}'\right\|^2 = 1$ y $\vec{x}'\perp\left[\left[\vec{e}_1\right]\right]$. Un argumento de homogeneidad, similar al usado anteriormente muestra que, en efecto, Q_2' es negativa para todo $\vec{x}$ de la forma $(0, x_2', ..., x_n')$.

Probaremos ahora que esto implica que los coeficientes b_{2j} y b_{j2} son cero para j$\neq$ 2. Es claro que $Q_2'\left[\vec{e}_2\right] = \lambda_2$ y por ende $b_{22} = \lambda_2$, así

$$Q_2'\left[\vec{e}_2 + \varepsilon\vec{e}_j\right] = b_{22}+b_{2j}\varepsilon+b_{j2}\varepsilon+b_{jj}\varepsilon^2-\lambda_2-\lambda_2\varepsilon^2 = 2b_{2j}\varepsilon+(b_{jj}-\lambda_2)\varepsilon^2 \leq 0$$

para todo real ε, lo que obliga a que $b_{2j} = b_{j2} = 0$ si $j \neq 2$. Por lo tanto

$$Q'\left[\vec{x}'\right] = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \sum_{i=3}^{n}\sum_{j=3}^{n} b_{ij}x_i'x_j'$$

Siguiendo esta linea de argumentación concluimos, inductivamente que, en efecto

$$Q\left[R\vec{x}'\right] = Q'\left[\vec{x}'\right] = \lambda_1\vec{x}_1'^2 + ... + \lambda_n x_n'^2$$

□

Observación 7. La demostración que hemos dado no es del todo completa puesto que hemos asumido que existia un vector $\vec{u}_j = (u_{1j}, ..., u_{nj})$ tal que $\left\|\vec{u}_j\right\|^2 = 1$, $\vec{u}_j\perp\left[\left[\vec{u}_1, ..., \vec{u}_{j-1}\right]\right]$ y

$$Q\left[u_{1j}, ..., u_{nj}\right] = Max\left\{Q\left[\vec{x}\right] : \left\|\vec{x}\right\|^2 = 1 \text{ y } \vec{x}\perp\left[\left[\vec{u}_1, ..., \vec{u}_{j-1}\right]\right]\right\}$$

Afortunadamente, en el caso de n variables uno puede mostrar la existencia usando la compacidad de la esfera unitaria $\mathbf{S}^{n-1} = \left\{\vec{x} \in \mathbf{E}^n : \left\|\vec{x}\right\|^2 = 1\right\}$ y la continuidad de las formas cuadráticas (toda función continua en un compacto alcanza su máximo y su mínimo). Este hecho no es trivial y, de hecho, no es válida en espacios de dimensión infinita. (En espacios de dimensión infinita la esfera unitaria no es compacta y no cualquier forma cuadratica es continua.)

3.3. Teorema Espectral para formas Bilineales y Matrices Simétricas

En la sección precedente mostramos que una forma cuadrática Q puede escribirse de manera única como

$$Q\left[\vec{x}\right] = Q\left[x_1, ..., x_n\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \text{ con } a_{ij} = a_{ji}$$

Una consecuencia inmediata de esto es que a cada forma cuadrática en n variables le podemos asociar de manera única la matriz $n \times n$ $A = (a_{ij})$. Es claro que A tiene la propiedad de que $a_{ij} = a_{ji}$.Con base en esto diremos que una matriz $A = (a_{ij})$ es simétrica si y sólo si $a_{ij} = a_{ji}$.para $i, j = 1, ..., n$.

La relación entre formas cuadráticas y matrices simétricas va mucho más allá del simple hecho de dar una manera conveniente de agrupar coeficientes.

Teorema 3.3.1. *A cada forma cuadrática Q le corresponde una única matriz simétrica A. Más aún*

$$Q\left[\vec{x}\right] = \langle A\vec{x}, \vec{x}\rangle$$

Esto es, la forma cuadrática se puede recuperar de la matriz A mediante el producto inteior.

Definición 3.3.2. Una **forma bilineal** en un espacio vectorial real V es una función $b : V \times V \to \mathbb{R}$ tal que

$$b(\alpha\vec{x} + \beta\vec{x}, \vec{z}) = \alpha b(\vec{x}, \vec{z}) + \beta b(\vec{y}, \vec{z}) \text{ y } b(\vec{x}, \alpha\vec{y} + \beta\vec{z}) = \alpha b(\vec{x}, \vec{y}) + \beta b(\vec{x}, \vec{z})$$

Si ademas $b(\vec{x}, \vec{y}) = b(\vec{y}, \vec{x})$ diremos que b es una **forma bilineal simétrica**.

De manera implícita, en el estudio de formas cuadráticas aparecen las formas bilineales, de hecho cada forma cuadrática da origen a una forma bilineal simétrica y con cada forma bilineal uno tiene una forma cuadrática. Discutimos ahora la manera en que estos dos coceptos se relacionan.

Es claro que dada la matriz A, la función $b(\vec{x}, \vec{y}) = \langle A\vec{x}, \vec{y}\rangle$ es una forma bilineal. Así pues, el teorma 3.3.1 nos permite asociar con cada forma cuadrática una forma bilineal de la siguiente manera: Dada una forma cuadrática Q, sea A la matriz asociada, entonces, la forma bilineal asociada a Q está dada por $b(\vec{x}, \vec{y}) = \langle A\vec{x}, \vec{y}\rangle$. Es claro que esta asociación es única. Más aún, puesto que $a_{ij} = a_{ji}$, $A^* = A$ y por ende

$$b(\vec{x}, \vec{y}) = \langle A\vec{x}, \vec{y}\rangle = \langle \vec{x}, A^*\vec{y}\rangle = \langle A\vec{y}, \vec{x}\rangle = b(\vec{y}, \vec{x})$$

Esto es, b es una forma bilineal simétrica. Además, uno puede obtener b directamente de Q

$$b(\vec{x},\vec{y}) = \frac{1}{4}(Q\left[\vec{x}+\vec{y}\right] - Q\left[\vec{x}-\vec{y}\right])$$

Esta última identidad puede obtenerse usando el hecho de que $Q\left[\vec{x}\right] = \langle A\vec{x},\vec{x}\rangle$, junto con la igualdad $b(\vec{x},\vec{y}) = \langle A\vec{x},\vec{y}\rangle = \langle A\vec{y},\vec{x}\rangle$.

TEOREMA 3.3.3. *A cada forma cuadrática Q le corresponde una única forma bilineal b. Más aún, dicha forma bilineal puede obtenerse de Q mediante la fórmula*

$$b(\vec{x},\vec{y}) = \frac{1}{4}(Q\left[\vec{x}+\vec{y}\right] - Q\left[\vec{x}-\vec{y}\right])$$

Recíprocamente, A toda forma bilineal simètrica $b\left(\vec{x},\vec{y}\right)$ le corresponde una única forma cuadrática, a saber

$$Q\left[\vec{x}\right] = b\left(\vec{x},\vec{x}\right)$$

Como mencionamos anteriormente una matriz B define una forma bilineal b mediante la relación

$$b(\vec{x},\vec{y}) = \langle B\vec{x},\vec{y}\rangle \tag{3.8}$$

Mostraremos ahora que a toda forma bilineal b le corresponde una única matriz B tal que la relación 3.8 es válida para todo $\vec{x}$ y $\vec{y}$ en E^n.

TEOREMA 3.3.4. *A toda forma bilineal simétrica b en n variables le corresponde una única matriz simétrica $A = (a_{ij})$ tal que*

$$b(\vec{x},\vec{y}) = \langle A\vec{x},\vec{y}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j$$

Además, los coeficientes a_{ij} se pueden obtener mediante la fórmula $a_{ij} = a(\vec{e}_i,\vec{e}_j)$.

DEMOSTRACIÓN. Dada la forma bilineal b definimos $a_{ij} = b(\vec{e}_i,\vec{e}_j)$. Es claro que si $\vec{x} = (x_1,\ldots,x_n)$ y $\vec{y} = (y_1,\ldots,\vec{y}_n)$, entonces, $\vec{x} = \sum_{i=1}^{n} x_i\vec{e}_i$ y $\vec{y} = \sum_{i=1}^{n} y_i\vec{e}_i$. Así, por bilinealidad

$$\begin{aligned} b(\vec{x},\vec{y}) &= b\left(\sum_{i=1}^{n} x_i\vec{e}_i, \sum_{j=1}^{n} y_j\vec{e}_j\right) = \sum_{i=1}^{n} x_i b\left(\vec{e}_i, \sum_{i=1}^{n} y_j\vec{e}_j\right) \\ &= \sum x_i\left\{\sum_{j=1}^{n} y_j b(\vec{e}_i,\vec{e}_j)\right\} = \sum_{i=1}^{n}\sum_{j=1}^{n} b(\vec{e}_i,\vec{e}_j)x_iy_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \end{aligned}$$

Por otra parte, si A es la matriz $A = (a_{ij})$, no es difícil mostrar que

$$\langle A\vec{x}, \vec{y}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j$$

de donde se sigue elresultado. La unicidad es inmediata. □

De este teorema se sigue que si b es una forma bilineal y $Q\left[\vec{x}\right] = b(\vec{x}, \vec{x})$, entonces

$$b(\vec{x}, \vec{y}) = \frac{1}{4}(Q\left[\vec{x}+\vec{y}\right] - Q\left[\vec{x}-\vec{y}\right])$$

Esto es completamente falso si b no es simétrica; la expresión del lado derecho de la igualdad siempre es simétrica en $\vec{x}$ y $\vec{y}$, lo que obliga a que b sea simétrica. Así pues uno puede recuperar la forma bilineal a partir de la forma cuadrática *solamente cuando la forma bilineal sea simétrica.* En positivo, podemos afirmar lo siguiente.

Teorema 3.3.5. *Existe una relación biunívoca entre formas cuadráticas, matrices simétricas y formas bilineales simétricas, la cual está dada por las identidades*

$$\begin{aligned} Q\left[\vec{x}\right] &= \langle A\vec{x}, \vec{x}\rangle = b(\vec{x}, \vec{x}) \\ b(\vec{x}, \vec{y}) &= \langle A\vec{x}, \vec{y}\rangle = \frac{1}{4}(Q\left[\vec{x}+\vec{y}\right] - Q\left[\vec{x}-\vec{y}\right]) \end{aligned}$$

De acuerdo con este resultado podemos reformular el teorema 3.2.4 en terminos de formas bilineales.

Teorema 3.3.6. *(Teorema Espectral versión **II**): Toda forma bilineal simétrica*

$$b(\vec{x}, \vec{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_i y_j$$

puede reducirse al la forma

$$b'(\vec{x}', \vec{y}') = \sum_{i=1}^{n} \lambda_i x_i' y_j'$$

mediante un cambio de coordenadas ortogonal. Esto es, existe una metriz ortogonal R tal que

$$b(R\vec{x}', R\vec{y}') = b'(\vec{x}', \vec{y}') = \sum_{i=1}^{n} \lambda_i x_i' y_j'$$

DEMOSTRACIÓN. Consideremos la forma cuadrártica $Q\left[\vec{x}\right] = b(\vec{x}, \vec{x})$. De acuerdo con el teorema 3.2.4 existe un matriz ortogonal R tel que

$$Q[R\vec{x}] = \sum_{i=1}^{n} \lambda_i {x'_i}^2$$

al aplicar el teorema anterior obtenemos que

$$\begin{aligned} b'(\vec{x}', \vec{y}') = b(R\vec{x}', R\vec{y}') &= \frac{1}{4}\left\{Q[R\vec{x}' + R\vec{y}'] - Q[R\vec{x}' - R\vec{y}']\right\} \\ &= \frac{1}{4}\left\{Q\left[R(\vec{x}+\vec{y})\right] - Q\left[R(\vec{x}-\vec{y})\right]\right\} \\ &= \frac{1}{4}\left\{\sum_{i=1}^{n} \lambda_i (x'_i + y'_i)^2 - \sum_{i=1}^{n} \lambda_i (x'_i - y'_i)^2\right\} = \sum_{i=1}^{n} \lambda_i x'_i y'_i \end{aligned}$$

□

Partiendo del teorema 3.3.5 tambien es posible reformular el Teorema Espectral en términos de matrices simétricas. En efecto si A es una matriz simétrica, entonces $b(\vec{x}, \vec{y}) = \langle A\vec{x}, \vec{y}\rangle$ es una forma bilineal simétrica y, de acuerdo al teorema anterior, existe una matriz ortogonal R tal que

$$\langle AR\vec{x}', R\vec{y}'\rangle = b(R\vec{x}', R\vec{y}') = \sum_{i=1}^{n} \lambda_i x'_i y'_j$$

Por otra parte es claro que $\sum_{i=1}^{n} \lambda_i x'_i y'_j = \langle \Lambda\vec{x}', \vec{y}'\rangle$, donde Λ es la matriz diagonal

$$\Lambda = \begin{pmatrix} \lambda_i & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Así para cualesquiera $\vec{x}$ y $\vec{y}$ en $\mathbf{E}^n$, $\langle AR\vec{x}', R\vec{y}'\rangle = \langle \Lambda\vec{x}', \vec{y}'\rangle$ con Λ una matriz diagonal.

Ahola bien, para poner este resultado en terminos exclusivamente de matrices recordemos que $\langle AR\vec{x}', R\vec{y}'\rangle = \langle R^*AR\vec{x}', \vec{y}'\rangle$, por lo tanto $\langle \Lambda\vec{x}', \vec{y}'\rangle = \langle R^*AR\vec{x}', \vec{y}'\rangle$ para todo $\vec{x}'$ y $\vec{y}'$, de aqui se sigue que R^*AR es justamente la matriz Λ. Esto demustra el siguiente teorema.

TEOREMA 3.3.7 (Teorema Espectral versión **III**). *Dada una matriz simétrica A existe una matriz ortogonal R tal que $R^*AR = \Lambda$ es una matrz diagonal.*

Como consecuencia inmediata de este resultado y del hecho de que para toda matriz ortogonal R, $R^* = R^{-1}$ tenemos las siguientes reformulaciones:

COROLARIO 3.3.8. *Dada una matriz simétrica A existe una matriz ortogonal R y una matriz diagonal Λ tales que*
i) $R^{-1}AR = \Lambda$
ii) $A = R\Lambda R^{*}$

La matriz R que "diagonaliza" a una matriz simétrica A se puede obtener aplicando el procedimiento de la sección 3.2 a la forma cuadrática asociada $Q\left[\vec{x}\right] = \langle A\vec{x}, \vec{x}\rangle$. Sin embargo, uno también puede caracterizar los vectores columna de R en términos de la matriz A.

Como se señalo en la sección 3.2 los vectores columna de R son los puntos críticos de $Q\left[x_1, \ldots, x_n\right]$ con la restricción $G\left[\vec{x}\right] = \|\vec{x}\|^2 = x_1^2 + \cdots + x_n^2 = 1$. De acuerdo con la teoría de multiplicadores de Lagrange, los puntos críticos, $\vec{u}_i = (u_{1i}, \cdots, u_{ni})$, deben cumplir las siguientes condiciones:

$$\begin{cases} \nabla Q\left[\vec{u}_i\right] = \lambda_i \nabla G\left[\vec{u}_i\right] \\ \|\vec{u}_i\|^2 = 1 \end{cases}$$

Ahora bien, no es difícil mostrar que si $Q\left[\vec{x}\right] = \langle A\vec{x}, \vec{x}\rangle$, entonces, $\nabla Q\left[\vec{x}\right] = 2A\vec{x}$. Por otra parte $\nabla G\left[\vec{x}\right] = 2\vec{x}$. Así que cada vector columna de la matriz ortogonal R debe satisfacer una ecuación de la forma

$$A\vec{u}_i = \lambda_i \vec{u}_i \tag{3.9}$$

para algún λ_i real.

TEOREMA 3.3.9. *Sea A una matriz simétrica. Entonces, si R es una matriz ortogonal que diagonaliza a A, cada uno de los vectores columna de R satisface una ecuación de la forma*

$$A\vec{u} = \lambda\vec{u}$$

Terminaremos señalando que, vía el teorema 3.3.5, las tres versiones del Teorema Espectral son equivalentes, en el sentido de que uno puede demostrar cualquiera dos versiones a partir de la otra. Más adelante discutiremos otras formulaciones de este mismo teorema, por supuesto, todas serán equivalentes.

3.4. Ecuaciones y el Teorema Espectral

En esta sección dicutiremos como se relaciona el Teorema Espectral con dos de los problemas fundamentales de la Matemática, a saber:

I) El problema de resolver *ecuaciones lineales*, esto es, encontrar el vector $\vec{u}$ tal que, para una vector $\vec{v}$ dado, satisfaga la igualdad

$$T\vec{u} = \vec{b}$$

donde T es un operador lineal.

II) El problema de resolver *ecuaciones lineales de evolución*: Encontrar la trayectoria $u(t)$ tal que

$$\begin{cases} \frac{du}{dt} = Tu \\ u(0) = u_0 \end{cases}$$

Donde u_0 es un vector fijo dado y T es un operador lineal.

El caso más sencillo de problemas del tipo (I) es el de un sistema de ecuaciones

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= y_n \end{aligned}$$

Este sistema se puede escribir en términos de matrices mediante la ecuación

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

donde A es la matriz $A = (a_{ij})$ y $\vec{x}$, $\vec{y}$ son los vectores columna $\vec{x} = (x_1, ..., x_n)$ y $\vec{y} = (y_1, ..., y_n)$. La base de la mayoria de los métodos usados para resolver este tipo de ecuaciones consiste en buscar un modo de transformar el sistema original en uno más *simple*. Aqui, el término *simple* debe entenderse en el sentido de que el nuevo sistema tenga la mayor cantidad posible de coeficientes iguales a cero; por supuesto el caso ideal sería cuando el sistema puede reducirse a la forma

$$\begin{aligned} \lambda_1 x'_1 &= y'_1 \\ \lambda_2 x'_2 &= y'_2 \\ &\vdots \\ \lambda_n x'_n &= y'_n \end{aligned} \tag{3.10}$$

En terminos de matrices esto quiere decir que A puede ser transformada en una matriz diagonal Λ.

En general, la tarea de encontrar el modo de transformar una matriz cualquiera en una matriz diagonal es harto difícil (y en muchos casos prácticamente imposible), sin embargo, en el caso de matrices simétricas, el teorema espectral nos permite reducir considerablemente el trabajo.

Del Teorema Espectral sabemos que existe una matriz ortogonal R tal que $R^*AR = \Lambda$ es uma matriz diagonal. Como $R^* = R^{-1}$ se tiene que dado $\vec{y}$, el vector $\vec{y}'$ tal que $\vec{y} = R\vec{y}'$ esta dado por $\vec{y}' = R^*\vec{y}$. Así, para resolver el sistema $A\vec{x} = \vec{y}$ nosotros consideramos el sistema $\Lambda\vec{x}' = \vec{y}'$.

Este sistema es de la forma 3.10, por tanto la solución $\vec{x}'$ (si $\lambda_i \neq 0$ para toda i) tiene como componentes

$$x_1' = \frac{y_1'}{\lambda_1}, x_2' = \frac{y_2'}{\lambda_2}, \dots, x_n' = \frac{y_n'}{\lambda_n}$$

Por otra parte

$$R^* A R \vec{x}' = \Lambda \vec{x}' = \vec{y}' = R^* \vec{y}$$

y utilizando nuevamenete el hecho de que $R^* = R^{-1}$ $(RR^* = I)$ concluimos que

$$AR\vec{x}' = \vec{y}$$

lo que uestra que $\vec{x} = R\vec{x}'$ es la solución del problema original.

Consideremos ahora el segundo problema en el caso en que T está dado por una matriz $A = (a_{ij})$, esto es, consideremos el sistema de ecuaciones diferenciables

$$\begin{aligned}
\dot{x}_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \cdots + a_{1n}x_n(t) \\
\dot{x}_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \cdots + a_{2n}x_n(t) \\
&\vdots \\
\dot{x}_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \cdots + a_{nn}x_n(t)
\end{aligned}$$

("·" denota derivada con respecto a t), junto con las condiciones iniciales

$$x_1(0) = x_0, x_2(0) = x_{02}, ..., x_n(0) = x_{0n}$$

En otros terminos queremos encontrar la trayecctoria $\vec{x}(t)$ que satisface el problema de condiciones iniciales

$$\begin{cases} \dot{\vec{x}}(t) = A\vec{x} \\ \vec{x}(t) = \vec{x}_0 \end{cases}$$

Como se sabe la solución a este problema está dada por

$$\vec{x}(t) = \mathbf{e}^{tA} x_0$$

donde $\mathbf{e}^{tA}$ es la matriz

$$\mathbf{e}^{tA} = \sum_{i=1}^{\infty} \frac{1}{k!}(tA)^k = \sum_{i=1}^{\infty} \frac{1}{k!} t^k A^k$$

Por supuesto el cálculo de esta matriz no es encillo, sin embargo, cuando la matriz A es simétrica el teorema espectral permite simpleficar considerablemente los cálculos. En efecto, si A es simétruca se tiene que existen una matriz ortogonal R y una matriz diagonal Λ tales que $A = R\Lambda R^*$ y como $R^* = R^{-1}$

$$\begin{aligned}
A^2 &= R\Lambda R^* R \Lambda R^* = R\Lambda^2 R^* \\
A^3 &= A^2 A = R\Lambda^2 R^* R \Lambda R^* = R\Lambda^3 R^*
\end{aligned}$$

en general uno tiene inductivamente, que

$$A^k = R\Lambda^k R^*$$

Usando este hecho y la continuidad de los operadores lineales en espacio de dimensión finita uno puede demostrar que

$$\mathbf{e}^{tA} = \sum_{i=1}^{\infty} \frac{1}{k!} t^k A^k = R\left(\sum_{i=1}^{\infty} \frac{1}{k!} t^k \Lambda^k\right) R^* = R\mathbf{e}^{t\Lambda} R^*$$

Ahora bien como Λ es una matriz diagonal

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

se tiene que

$$\Lambda^k = \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}$$

y por ende

$$\sum_{i=1}^{\infty} \frac{1}{k!} t^k \Lambda^k = \begin{bmatrix} \sum_{i=1}^{\infty} \frac{1}{k!} t^k \lambda_1^k & 0 & \dots & 0 \\ 0 & \sum_{i=1}^{\infty} \frac{1}{k!} t^k \lambda_2^k & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sum_{i=1}^{\infty} \frac{1}{k!} t^k \lambda_n^k \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{e}^{t\lambda_1} & 0 & \dots & 0 \\ 0 & \mathbf{e}^{t\lambda_2} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{e}^{t\lambda_n} \end{bmatrix}$$

y por lo tanto

$$\mathbf{e}^{t\Lambda} = \begin{bmatrix} \mathbf{e}^{t\lambda_1} & 0 & \dots & 0 \\ 0 & \mathbf{e}^{t\lambda_2} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{e}^{t\lambda_n} \end{bmatrix}$$

Una vez calculada la matriz $\mathbf{e}^{t\Lambda}$, la solución de nuestro problema de valores iniciales se puede calcular mediante un simple producto de matrices, a saber

$$\vec{x}(t) = R\mathbf{e}^{t\Lambda}R^*\vec{x}_0$$

Terminamos señalando que en muchos casos no es necesario calcular explícitamente las soluciones ya que gran parte de la información acerca del comportamiento de las soluciones está determinado directamente por la matriz $\mathbf{e}^{t\Lambda}$.

3.5. Vectores y Valores Propios. Subespacios Invariantes

En esta sección presentaremos el teorema espectral desde un enfoque un tanto diferente. El problema que nos servira de motivación para desarrollar las ideas de esta sección podría plantearse de la siguiente manera: Sabemos que dado un operador lineal $T : V \to V$, (donde V es un espacio de dimensión finita con producto interior) se puede representar como una matriz A y que esta representación depende de la base de V que se elija. El problema es *?'Bajo qué condiciones es posible dar una representación matricial lo más sencilla posible de un operador?*

Dos puntos de esta formulación requieren ser puntualizados; por lo *más sencilla posible* entendemos, por supuesto, una matriz diagonal. El segundo punto tiene que ver con buscar explotar la estructura determinada por el producto interior de V, en este sentido el término *representación matricial* se entiende como la representación de nuestro operador en una base *ortonormal* de V. En este contexto nuestro problema queda formulado como sigue:

¿Bajo qué condiciones la matriz asociada a un operador $T : V \to V$ en una base ortonormal es una matriz diagonal?

Una de las claves es el teorema 3.3.9. En efecto, si es posible encontrar una base de vectores que cumplan una ecuación de la forma $T\vec{u} = \lambda\vec{u}$ para diferentes valores de λ, la matriz asociada a T en dicha base será una matriz diagonal, esto da origen a nuestra siguiente definición

Definición 3.5.1. Dado un operador lineal $T : V \to V$ diremos que $\vec{u} \in V$ es un **vector propio o eigenvector** de T si y sólo si

i) $\vec{u} \neq \vec{0}$

ii) Existe λ tal que

$$T\vec{u} = \lambda\vec{u}$$

De manera análoga si λ es un número para el cual existe un vector $\vec{u} \neq 0$ tal que $T\vec{u} = \lambda\vec{u}$ diremos que λ es un **valor propio o eigenvalor** de T.

Si $\vec{u}$ es un vector propio y λ es el número tal que $T\vec{u} = \lambda\vec{u}$ diremos que λ es el valor propio asociado a $\vec{u}$. Recíprocamente, dado un valor propio λ, un vector $\vec{u} \neq \vec{0}$ para el cual $T\vec{u} = \lambda\vec{u}$ es llamado un vector propio asociado al eigenvector λ.

El otro elemento clave es en el análisis de nuestro problema es el de subespacio invariante.

DEFINICIÓN 3.5.2. Dado un operador lineal $T : V \to V$ diremos que un subespacio E de V es un **subespacio invariante** de T si para todo $\vec{u} \in E$, $T\vec{u} \in E$, esto es si $TE \subset E$.

Con este concepto a nuestro alcance uno puede ver que una condición "natural" para poder obtener una representación sencilla de T es que T tenga la propiedad de que si E es un subespacio invariante de T entonces $E^{\perp}$ también sea un subespacio invariante de T. En efecto, bajo estas condiciones se puede tomar una base ortonormal de E y completarla con una base ortonornal de $E^{\perp}$ para formar una base ortonormal de V. En términos de esta base ortonormal de V, uno puede verificar sin mucha dificultad que la matriz asociada a T es de la forma

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

donde A_1 y A_2 son matrices cuadradas de dimensión menor que la de A.

La relación entre subespacios invariantes y vectores y valores propios está dada por el siguiente resultado:

TEOREMA 3.5.3. *Sea $T : V \to V$ un operador lineal y sea λ un valor propio de T. Entonces, el conjunto*

$$E_\lambda = \{\vec{u} \in V : T\vec{u} = \lambda\vec{u}\}$$

es un subespacio invariante de T. Y el conjunto de vectores propios asociados a un valor mismo propio de T forman un subespacio invariante de T.

La demostración de este hecho es consecuencia inmediata de la linealidad y se deja como ejercicio.

El siguiente resultado da una respuesta a cuando un operador tiene la propiedad de que si $TE \subset E$, entonces, $TE^{\perp} \subset E^{\perp}$. En esta dirección uno tiene el siguiente resultado.

LEMA. *: Si $T : V \to V$ es un operador lineal y E es un subespacio invariante de T, entonces, $E^{\perp}$ es un subespacio invariante de T^*. En particular, si T es un operador auto-adjunto, i.e., si $T = T^*$, entonces, $E^{\perp}$ es un subespacio invariante de T si y sólo si E es un subespacio invariante de T.*

Demostración. Puesto que para toda $\vec{u} \in E$ se tiene que $T\vec{u} \in E$ y por lo tanto

$$\langle T\vec{u}, \vec{v}\rangle = 0 \text{ para todo } \vec{v} \in E$$

Como $\langle T\vec{u}, \vec{v}\rangle = \langle \vec{u}, T^*\vec{v}\rangle$ se tiene que para todo $\vec{u} \in E$ y $\vec{v} \in E^{\perp}$

$$\langle \vec{u}, T^*\vec{v}\rangle = 0$$

lo que muestra que $T^*\vec{v} \in E^{\perp}$ para todo $\vec{v} \in E^{\perp}$. □

El lema muestra que para que un operador se pueda representar por una matriz diagonal en términos de una base ortonormal, debe existir alguna relación entre éste y su adjunto.

Como señalamos anteriormente la existencia de espacios invariantes permite encontrar representaciones matriciales más simples de un operador. Por supuesto, entre más subespacios invariantes mutuamente perpendiculares tenga un operador, su representación matricial en términos de la base ortonormal adecuada será más sencilla.

El teorema 3.5.3 nos dice que para encontrar subespacios invariantes de un operador debemos encontrar sus valores propios y los correspondientes vectores propios, es decir, debemos determinar para que números λ la ecuación $T\vec{u} = \lambda\vec{u}$ tiene soluciones diferentes de cero y cuáles son las dichas soluciones. Puesto que $T\vec{u} = \lambda\vec{u}$ tiene solución si y sólo si $T\vec{u} - \lambda\vec{u} = 0$, este problema es equivalente al problema de determinar bajo que condiciones la ecuación

$$(T - \lambda I)\vec{u} = 0 \tag{3.11}$$

tiene soluciones diferentes de cero y cúales son dichas soluciones. En otros terminos, para que valores de λ el $\ker(T - \lambda I) \neq \{0\}$. Note que $\ker(T - \lambda I) = \{\vec{u} \in V : T\vec{u} = \lambda\vec{u}\} = E_{\lambda}$.

De acuerdo con la discusión del capítulo anterior $\ker(T - \lambda I) \neq \{0\}$ si y sólo si $\det(T - \lambda I) = 0$. Así, λ es un valor propio de Tsi y sólo si λ es solución de la ecuación $\det(T - \lambda I) = 0$. Esta ecuación es una ecuación polinomial en λ de grado igual a la dimensión de V.

Teorema 3.5.4. *: Sea V un espacio vectorial de dimensión n y sea $T : V \to V$ un operador lineal. Entonces, existen a lo más n valores propios, $\lambda_1, \ldots, \lambda_n$ de T (contando multiplicadores). Además, para cada λ_i el subespacio $E_{\lambda_i} = \ker(T - \lambda_i I)$ es un subespacio invariante de T.*

Cada subespacio E_{λ_i} es llamado el *subespacio propio* o *eigenespacio* asociado al valor propio λ_i y por supuesto todo $\vec{u} \in E_{\lambda_i}$ es un vector propio de T con eigenvalor λ_i y esto es, $T\vec{u} = \lambda_i\vec{u}$.

En general, como acabamos de mencionar, los valores propios no tienen por que ser reales y tampoco los espacios propios tienen por que ser mutuamente perpendiculares, sin embargo, en el caso de operadores auto-adjunto uno tiene el siguiente resultado.

PROPOSICIÓN 3.5.5. *: Sea $T : V \to V$ un operador auto-adjunto (i.e $T = T^*$). Entonces, si λ_i y λ_j son valores propios de T y $\lambda_i \neq \lambda_j$ entonces E_{λ_i} es ortogonal a E_{λ_j}, esto es,*

$$\langle \vec{v}, \vec{w} \rangle = 0 \; \forall \vec{v} \in E_{\lambda_i} \text{ y } \forall \vec{w} \in E_{\lambda_j}$$

DEMOSTRACIÓN. Sea $\vec{v} \in E_{\lambda_i}$ y sea $\vec{w} \in E_{\lambda_j}$. Puesto que T es auto-adjunto

$$\langle T\vec{v}, \vec{w} \rangle = \langle \vec{v}, T\vec{w} \rangle$$

Por otro lado sabemos que $T\vec{v} = \lambda_i \vec{v}$ y $T\vec{w} = \lambda_j \vec{w}$ y por ende

$$\langle T\vec{v}, \vec{w} \rangle = \langle \lambda_i \vec{v}, \vec{w} \rangle = \lambda_i \langle \vec{v}, \vec{w} \rangle \;\text{ y }\; \langle \vec{v}, T\vec{w} \rangle = \langle \vec{v}, \lambda_j \vec{w} \rangle = \lambda_j \langle \vec{v}, \vec{w} \rangle$$

en consecuencia $\lambda_i \langle \vec{v}, \vec{w} \rangle = \lambda_j \langle \vec{v}, \vec{w} \rangle$, esto es

$$\left(\lambda_i - \lambda_j\right) \langle \vec{v}, \vec{w} \rangle = 0$$

y como $\lambda_i \neq \lambda_j$ necesariamente $\langle \vec{v}, \vec{w} \rangle = 0$, lo que muestra que $E_{\lambda_i} E_{\lambda_j}$. □

TEOREMA 3.5.6. *(Teorema Espectral versión IV): Sea $T : V \to V$ un operador auto-adjunto. Entonces existe una base ortonormal de vectores propios de T. Y en terminos de esta base, la matriz asociada a T es una matriz diagonal.*

DEMOSTRACIÓN. La demostarción por inducción sobre la dimensión de V. Si $\dim V = 1$ el resultado es trivial. Supongamos que el teorema es válido para espacios de dimensión $n - 1$, Si $\dim V = n$, sabemos que los resultados 3.5.4 y 3.5.5 que al menos existe un valor propio real λ_1. Sea $\vec{u}_1$ un vector propio asociado a λ_1. Es claro que el subespacio $\left[\left[\vec{u}_1\right]\right]$, generao por $\vec{u}_1$, es un subespacio invariante de T y por el teorema 3.5 $\left[\left[\vec{u}_1\right]\right]^{\perp}$ también es un subespacio invariante de T, además la dimensión de $\left[\left[\vec{u}_1\right]\right]^{\perp}$ es $n - 1$. No es difícil ver que T restringido a $\left[\left[\vec{u}_1\right]\right]^{\perp}$, $T : \left[\left[\vec{u}_1\right]\right]^{\perp} \to \left[\left[\vec{u}_1\right]\right]^{\perp}$ es un operador auto-adjunto. Así, por hipótesis de inducción, existe una base ortonormal de $\left[\left[\vec{u}_1\right]\right]^{\perp}$ farmada por vectores propios de T, digamos $\{\vec{u}_1, ..., \vec{u}_n\}$. Puesto que $V = \left[\left[\vec{u}_1\right]\right] \oplus \left[\left[\vec{u}_1\right]\right]^{\perp}$ y se tiene que $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ es una base ortonormal de V y cada uno de los vectores de la base es un vector propio de T. Con este resultado no es dificil mostrar que la matriz asociada a T en esta base es la matriz

$$\Lambda = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$$

donde $\lambda_1, ..., \lambda_n$ son los valores propios de T asociados a $\vec{u}_1, ..., \vec{u}_n$, respectivamente. □

Observación 8. Este resultado no es sorprendente si se toma en cuenta los resultados de la sección 3.3 y el hecho de que los operadores auto-adjunto son las generalización natural de las matrices simétricas. Es más, uno puede dar una demostración de la versión IV a partir del teorema 2.2.4 y los resultados de las secciones 3.2 y 3.3. Esto se deja como ejercicio.

3.6. Expanciones del Tipo de Fourier y Ecuaciones

En esta seccion discutiremos nuevamente los problemas I) y II) plantedos en la sección 3.4, pero partiendo del hecho de que tenemos una base ortonormal de vectores proipios. Así pues, asumiremos que $T : V \to V$ es un operador auto-adjunto y que $\{\vec{u}_1, ..., \vec{u}_n\}$ es una base ortonormal de vectores propios de T.

El primer problema a considerar es el de la resolución de la ecuación lineal

$$T\vec{u} = \vec{v}$$

La idea es reformular el problema en términos de la base ortonormal de vectores propios. Empecemos por observar que como $\{\vec{u}_1, ..., \vec{u}_n\}$ es una base ortonormal, entonces, todo $\vec{u} \in V$ tiene una representación del tipo de Fuorier

$$\vec{u} = \langle \vec{u}, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{u}, \vec{u}_2 \rangle \vec{u}_2 + ... + \langle \vec{u}, \vec{u}_n \rangle \vec{u}_n = \sum_{i=1}^{n} \langle \vec{u}, \vec{u}_i \rangle \vec{u}_i \tag{3.12}$$

Ahora bien, como cada $\vec{u}_i$ es un vector propio de T con eigenvalor λ_i, se tiene que

$$T\vec{u} = \sum_{i=1}^{n} T\left(\langle \vec{u}, \vec{u}_i \rangle \vec{u}_i\right) = \sum_{i=1}^{n} \langle \vec{u}, \vec{u}_i \rangle T\vec{u}_i = \sum_{i=1}^{n} \langle \vec{u}, \vec{u}_i \rangle \lambda_i \vec{u}_i$$

Por otra parte, $\vec{v}$ también tiene una expanción del tipo de Fourier, digamos

$$\vec{v} = \sum_{i=1}^{n} \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i$$

De esto concluimos que, en términos de la base ortonormal de eigenvectores de T, la ecuación $T\vec{u} = \vec{v}$ se reducir a la ecuación

$$\sum_{i=1}^{n} \lambda_i \langle \vec{u}, \vec{u}_i \rangle \vec{u}_i = \sum_{i=1}^{n} \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i \tag{3.13}$$

Como $\{\vec{u}_1, ..., \vec{u}_n\}$ es un conjunto linealmente independiente, para que exista una solución de 3.13 es necesario y suficiente que

$$\lambda_i \langle \vec{u}, \vec{u}_i \rangle = \langle \vec{v}, \vec{u}_i \rangle, \ i = 1, ..., n$$

Por supueto, si $\lambda_i \neq 0$ para $i = 1, ..., n$, entonces $\langle \vec{u}, \vec{u}_i \rangle = \lambda_i^{-1} \langle \vec{v}, \vec{u}_i \rangle$ y al substituir en (3.12) se obtiene que

$$\vec{u} = \sum_{i=1}^{n} \lambda_i^{-1} \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i$$

es la solución a nuestro problema.

TEOREMA 3.6.1. *: Sea $T : V \to V$ un operador auto-adjunto. Sea $\{\vec{u}_1, ..., \vec{u}_n\}$ una base ortonormal de vectores propios de T. Para que $\vec{u} \in V$ sea solución de la ecuación $T\vec{u} = \vec{v}$ es necesario y suficiente que $\lambda_i \langle \vec{u}, \vec{u}_i \rangle = \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i$ $i = 1, ..., n$, donde λ_i es el valor propio asociado a $\vec{u}_i$. En particular, si todos los valores propios de T son direrentes de cero, la solución de la ecuación es el vector*

$$\vec{u} = \sum_{i=1}^{n} \lambda_i^{-1} \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i$$

Este resultado también permite determinar cuando T es invertible y calcular el operador inverso en términos de expanciones del tipo de Fourier.

COROLARIO 3.6.2. *Si $T : V \to V$ es un operador auto-adjunto. Entonces, T es invertible si y sólo si todos sus valores propios son diferentes de cero. En tal caso el operador inverso T^{-1} está dado por el operador*

$$T^{-1}v = \sum_{i=1}^{n} \lambda_i^{-1} \langle v, u_i \rangle u_i$$

Pasemos ahora al problema II, esto es, al problema de evolución

$$\begin{cases} \frac{du}{dt} = Tu \\ u(0) = u_0 \end{cases}$$

La solución a este vector es una curva $u(t)$ en V, la cual, para cada t, se puede escribir en términos de la base ortonormal de vectores propios en al forma

$$u(t) = \sum_{i=1}^{n} \langle u(t), u_i \rangle u_i \tag{3.14}$$

Como

$$Tu(t) = \sum_{i=1}^{n} \langle u(t), \vec{u}_i \rangle T\vec{u}_i = \sum_{i=1}^{n} \langle u(t), \vec{u}_i \rangle \lambda_i \vec{u}_i$$

y $\vec{u}_0 = \sum_{i=1}^{n} \langle \vec{u}_0, \vec{u}_i \rangle \vec{u}_i$, la ecuación de evolución en términos de la base de vectores propios toma la forma

$$\sum_{i=1}^{n} \frac{d \langle u(t), \vec{u}_i \rangle}{dt} \vec{u}_i = \sum_{i=1}^{n} \lambda_i \langle u(t), \vec{u}_i \rangle \vec{u}_i$$

con las condiciones iniciales

$$\vec{u}(0) = \sum_{i=1}^{n} \langle \vec{u}_0, \vec{u}_i \rangle \vec{u}_i$$

Nuevamente la independencia lineal de la base reduce nuestro problema a resolver n ecuaciones lineales de primer orden:

$$\begin{cases} \frac{d\langle u(t), \vec{u}_i \rangle}{dt} = \lambda_i \langle u(t), \vec{u}_i \rangle \\ \langle u(0), \vec{u}_i \rangle = u(t) = \langle \vec{u}_0, \vec{u}_i \rangle \end{cases} \quad i = 1, ..., n$$

La solución de este tipo de ecuaciones es bien conocida, son exponencial

$$\langle u(t), \vec{u}_i \rangle = \mathbf{e}^{\lambda_i t} \langle \vec{u}_0, \vec{u}_i \rangle$$

Al substituir estos valores en (3.14) obtenemos que la solución de la ecuación de evolución es

$$u(t) = \sum_{i=1}^{n} \mathbf{e}^{\lambda_i t} \langle \vec{u}_0, \vec{u}_i \rangle \vec{u}_i$$

Teorema 3.6.3. *Sea $T : V \to V$ un operador auto-adjunto y sea $\{\vec{u}_1, \ldots, \vec{u}_n\}$ una base ortonormal de vectores propios de T. Entonces, la solución de el problema de evolución*

$$\begin{cases} \frac{du}{dt} = Tu \\ u(0) = \vec{u}_0 \end{cases}$$

es la trayectoria $u(t) = \sum_{i=1}^{n} \mathbf{e}^{\lambda_i t} \langle \vec{u}_0, \vec{u}_i \rangle \vec{u}_i$, donde λ_i es el valor propio de T asociado a $\vec{u}_i$.

Por analogía con el caso unidimencional (la solución de $du/dt = au$ es $\mathbf{e}^{ta} u_0$) y con el caso de matrices simétricas (la solución de $du/dt = Au$ es $\mathbf{e}^{tA} \vec{u}_0$) uno define e^{tT} para un operador auto-adjunto T, como el operador

$$e^{tT} \vec{u} = \sum_{i=1}^{n} \mathbf{e}^{\lambda_i t} \langle \vec{u}, \vec{u}_i \rangle \vec{u}_i \; \forall \vec{u} \in V$$

donde $\{\vec{u}_1, ..., \vec{u}_n\}$ es una base ortonormal de vectores propios de T y $\lambda_1, \ldots, \lambda_n$ son los valores propios correspondientes.

3.7. Proyecciones Ortogonales y Teorema Espectral

En muchos casos resulta más conveniente formular las cosas en términos de operadores y no tan solo mediante matrices o representanciones del tipo de Fourier. Afortunadamente el teorema espectral también se puede formular directamente mediante operadores. Para ver como se llega a esto retomemos la discución de la sección 3.5 y en particular los resultados 3.5.4 y 3.5.5.

Empezaremor por demostrar que en el caso de operadores auto-adjuntos los subespacios propios "cubren" todo el espacio.

TEOREMA 3.7.1. *Sea $T : V \to V$ un operador auto-adjunto. Asúmase que $\lambda_1, \lambda_2, ..., \lambda_k$ son todos los valores propios **distintos** de T. Entonces, si $E_{\lambda_1}, E_{\lambda_2}, ..., E_{\lambda_k}$, son los subespacios propios asociados a dichos eigenvalores, se tiene que*

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus ... \oplus E_{\lambda_k} = \bigoplus_{i=0}^{k} E_{\lambda_i}$$

DEMOSTRACIÓN. De la proposición 3.5.5 sabemos que los subespacios E_{λ_i} son mutuamente perpendiculares. Así que sólo debemos demostrar que, en efecto, V es la suma de los subespacios invariantes; lo haremos por reducción al absurdo.

Denotemos por E a la suma directa de los E_{λ_i}; $E = \bigoplus_{i=0}^{k} E_{\lambda_i}$. Puesto que todo elemento $\vec{u} \in E$ es de la forma $\vec{u} = \vec{u}_1 + ... + \vec{u}_k$, con $\vec{u}_i \in E_{\lambda_i}$ y $T\vec{u}_i$ está en E_{λ_i} se tiene que E es un subespacio invariante de T. Como T es auto-adjunto, $E^{\perp}$ también es un subespacio invariente de T. Ahora bien, uno puede ver que T retringido a $E^{\perp}$ (esto es, $T : E^{\perp} \to E^{\perp}$) es un operador auto-adjunto. Luego, si suponemos que $E^{\perp} \neq \{0\}$, el operador retringido debe tener un vertor propio $\vec{v} \in E^{\perp}$, $\vec{v} \neq 0$, con eigenvalor λ. Pero $\vec{v}$ también sería un vector propio del operador T definido en todo V, por lo tanto λ tendria que ser igual a algún λ_i y $\vec{v} \in \ker(T - \lambda_i I) = E_{\lambda_i} \subset E$ lo que implica que $\vec{v} = 0$. Contradicción. Lo que muestra que necesariamente $E^{\perp} = \{0\}$ y por tanto $E = V$. □

Los subespacios propios E_{λ_i} no sólo tienen la ventaja de ser invariantes de T sino además la restricción de T a cualquier de estos subespacios es, simplemente, un operador de multiplicación por una constante, a saber el valor propio correspondiente. En particular, si P_{λ_i} es el operador de proyección sobre E_{λ_i}, entonces, para todo $\vec{u} \in V$

$$TP_{\lambda_i}\vec{u} = \lambda_i P_{\lambda_i}\vec{u}$$

Este hecho junto con los teoremas 2.3.9 y 3.7.1 se puede obtener una descomposición de T en términos de las proyecciones ortogonales sobre los subespacios propios. En efecto, del teorema 2.3.9 y 3.7.1 pueden ser explotadas para obtener una descomposición de T en términos de las proyecciones ortogonales sobre los subespacios propios. En efecto, del teorema 2.3.9 sabemos que $P = P_{\lambda_1} + P_{\lambda_2} + \ldots + P_{\lambda_k}$ es la proyección ortogonal sobre $\bigoplus_{i=0}^{k} E_{\lambda_i}$ y como $V = \bigoplus_{i=0}^{k} E_{\lambda_i}$ se tiene que

$$\vec{u} = P_{\lambda_1}\vec{u} + P_{\lambda_2}\vec{u} + \ldots + P_{\lambda_k}\vec{u} = \sum_{i=1}^{k} P_{\lambda_i}\vec{u}$$

en consecuencia

$$\begin{aligned} T\vec{u} &= TP_{\lambda_1}\vec{u} + TP_{\lambda_2}\vec{u} + \ldots + TP_{\lambda_k}\vec{u} = \vec{u} \\ &= \lambda_1 P_{\lambda_1}\vec{u} + \lambda_2 P_{\lambda_2}\vec{u} + \ldots + \lambda_k P_{\lambda_k}\vec{u} = \sum_{i=1}^{k} \lambda_i P_{\lambda_i}\vec{u} \end{aligned}$$

TEOREMA 3.7.2 (Teorema Espectral versión V). *Sea $T : V \to V$ un operador auto-adjunto y sean $\lambda_1, \ldots, \lambda_k$* ***todos*** *los valores propios* ***distintos*** *de T. Entonces, si P_{λ_i} son las proyecciones ortogonales sobre los eigenespacios E_{λ_i}, se tiene que*

$$T = \sum_{i=1}^{k} \lambda_i P_{\lambda_i}$$

A continuación demostraremos la versión análoga del corolario 3.6.2 y del teorema 3.7.1.

TEOREMA 3.7.3. *Bajo las hipótesis del teorema 3.7.1 uno tiene*

i) $T^m = \sum_{i=0}^{k} \lambda_i^m P_{\lambda_i}$

ii) $\mathbf{e}^{tT} = \sum_{i=0}^{k} \mathbf{e}^{t\lambda_i} P_{\lambda_i}$

iii) Si T es invertible (i.e. si todos los valores propios de T son diferentes de cero), entonces

$$T^{-1} = \sum_{i=1}^{k} \lambda_i^{-1} P_{\lambda_i}$$

DEMOSTRACIÓN. : Usando el hecho de que $P_{\lambda_i} \perp P_{\lambda_j}$ uno tiene que

$$T^2 = TT = \left(\sum_{i=1}^{k} \lambda_i P_{\lambda_i}\right)\left(\sum_{i=1}^{k} \lambda_i P_{\lambda_i}\right) = \sum_{i=1}^{k} \lambda_i^2 P_{\lambda_i}$$

de manera inductiva uno puede entonces concluir i).

Para demostrar ii) uno usa i):

$$\mathbf{e}^{tT} = \sum_{m=1}^{\infty} \frac{1}{m!} t^m T^m = \sum_{m=1}^{\infty} \frac{1}{m!} t^m \left(\sum_{i=1}^{k} \lambda_i^m P_{\lambda_i} \right)$$

$$= \sum_{m=1}^{\infty} \left(\sum_{i=1}^{k} \frac{1}{m!} t^m \lambda_i^m P_{\lambda_i} \right) = \sum_{i=0}^{k} \left(\sum_{m=1}^{\infty} \frac{1}{m!} t^m \lambda_i^m \right) P_{\lambda_i} = \sum_{i=1}^{k} \mathbf{e}^{t\lambda_i} P_{\lambda_i}$$

Finalmente para ver iii) explotamos nuevamente el que $P_{\lambda_i} \perp P$

$$\left(\sum_{i=1}^{k} \lambda_i P_{\lambda_i} \right) \left(\sum_{i=1}^{k} \lambda_i^{-1} P_{\lambda_i} \right) = \sum_{i=1}^{k} \lambda_i \lambda_i^{-1} P_{\lambda_i} = I$$

lo que muestra que $T^{-1} = \sum_{i=1}^{k} \lambda_i^{-1} P_{\lambda_i}$. □

El teorema sugiere de manera natural la siguiente definición:

Definición 3.7.4. Sea $T : V \to V$ un operador auto-adjunto. Para cada función f defina en los valores propios de T, definimos el operador $f(T)$ como el operador

$$f(T) = \sum_{i=1}^{k} f(\lambda_i) P_{\lambda_i}$$

Esta definición puede reformularse en términos de representaciones del tipo de Fourier si uno expresa a cada proyección mediante una base ortonormal de vectores propios de T.

Teorema 3.7.5. *: Sea $T : V \to V$ un operador auto-adjunto. Sea $\{\vec{u}_1, ..., \vec{u}_n\}$ una base ortonormal de vectores propios de T. Para cada función f con valores reales o complejos definida en los valores propios de T*

$$f(T)\vec{u} = \sum_{i=1}^{n} f(\lambda_i) \langle \vec{u}, \vec{u}_i \rangle \vec{u}_i$$

donde λ_i es el valor propio asociado a $\vec{u}_i$.

3.8. Problemas

Los problemas 1 a 4 son sobre espacios de *dimensión finita*.

1. Sea $T : V \to V$ un operador auto-adjunto y sea $Q[\vec{x}] = \langle T\vec{x}, \vec{x} \rangle$ la forma cuadrática asociada. Musetre que $T\vec{x} = \vec{y}$ tiene una única solución si y sólo si existe una constante $c > 0$ tal que $c \|\vec{x}\|^2 \leq Q\left[\vec{x}\right]$ (Sug.: $T\vec{x} = \vec{y} \Longleftrightarrow \langle T\vec{x}, \vec{u} \rangle = \langle \vec{y}, \vec{u} \rangle \, \forall \vec{u} \in V$ y use el problema 1 del Capítulo 2)

2. Muestre que un operador normal T es invertible si y sólo si todo valor propio de T es diferente de cero. Pruebe que en tal caso

a) T^* también es invertible y que $(T^*)^{-1} = (T^{-1})^*$

b) T^{-1} es normal

c) $\vec{u}$ es un vector propio de T^{-1} si y sólo si $\vec{u}$ es un vector propio de T y μ es un valor propio de T^{-1} si y sólo si $\frac{1}{\mu}$ es un valor propio de T.

3. ***a***) Encuentre los valores y vectores propios de la matriz

$$A = \begin{bmatrix} 3/2 & -1 & 0 \\ -1 & 1/2 & 3/\sqrt{2} \\ 0 & 3/\sqrt{2} & 3 \end{bmatrix}$$

b) Diga si A es o no invertible. En caso afirmativo encuentre la matriz A^{-1}.

c) Calcule $\sqrt{A}$, esto es, encuentre la matriz B tal que $B^2 = A$.

d) Encuentre las soluciones del sistema de ecuaciones diferenciables

$$\frac{d^2u}{dt^2} = Au$$

con condicones iniciales $u'(0) = (1,1,1)$ y $u(0) = (2,1,2)$. (Sug.: Calcule $\mathrm{e}^{t\sqrt{A}}$.)

4. Considere el operador L definido en el ejemplo 1.4 del Capítulo 1

a) Muestre que L es auto-adjunto.

b) Encuentre la base ortonormal de vectores propios.

c) Encuentre la solución $u(t,x)$ del problema de Cauchy para la ecuación de derivadas parciales

$$\begin{cases} \frac{\partial u}{\partial t} = (x^2-1)\frac{\partial^2 u}{\partial x^2} + 2x\frac{\partial u}{\partial x} \\ u(0,x) = x^3 - x^2 \end{cases}$$

(Sug.: La ecuación $\frac{\partial u}{\partial t} = (x^2-1)\frac{\partial^2 u}{\partial x^2} + 2x\frac{\partial u}{\partial x}$ puede verse como $\dot{u} = Lu$ y $u(0,x) = x^3 - x^2$ como el vector en $P_3[x]$ dado por (0,0,-1,1) en la base $\{1, x, x^2, x^3\}$).

5. Considere el espacio $L^2(\mathbb{R}, e^{-x^2}dx)$ de todas las funciones medibles tales que

$$\int_{-\infty}^{\infty} |\, u(x)\,|^2\, e^{-x^2} dx < \infty$$

a) Muetre que

$$\langle u, \vec{v} \rangle = \int_{-\infty}^{\infty} u(x)\vec{v}(x) e^{-x^2} dx < \infty$$

es un producto interior para $L^2(\mathbb{R}, e^{-x^2}dx)$.

b) Muestre que $L^2(\mathbb{R}, e^{-x^2}dx)$ con este producto interior es un spacio de Hilbert.

c) Muestre que todo polinomio esta en $L^2(\mathbb{R}, e^{-x^2}dx)$. (Sug.: Muestre que x^n esta en $L^2(\mathbb{R}, e^{-x^2}dx)$.)

d) Aplique Gram-Schmidt al conjunto $\{1, x, x^2, \ldots\}$ para obtener la sucesión de polinomios

$$\frac{1}{(2^n n!\sqrt{\pi})^{\frac{1}{2}}} H_n(x)$$

y demuestre que

$$H_0(x) = 1 \text{ y } H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left[e^{-x^2}\right]$$

Los polinomios H_n son llamados polinomios de *Hermite*.

e) Muestre que $dH_n/dx = 2nH_{n-1}$. Use esta identidad para mostrar directamente que los polinomios de Hermite son ortogonales con respecto al producto interior definido en el inciso a). Calcule su norma.

f) Muestre que los polinomios de Hermite son funciones propias del operador

$$Hu = \frac{d^2u}{dx^2} - 2x\frac{du}{dx}$$

6. Sea $\mathbf{H}$ un espacio de Hilbert y sea $\mathbf{H}^*$ su dual. Defina el producto $\langle\cdot,\cdot\rangle_*$ en $\mathbf{H}^*$ como

$$\langle f,g\rangle_* = \langle \vec{v}_f, \vec{v}_g\rangle \quad \forall\, f,g \in \mathbf{H}^*$$

donde $\vec{v}_f$ y $\vec{v}_g$ son vectores en $\mathbf{H}$ asociados a f y g por el teorema de Riesz.

a) Muestre que $\langle\cdot,\cdot\rangle_*$ es un producto interior en $\mathbf{H}^*$ y que la norma de una funcional, $\|\cdot\|_*$, es igual a la norma inducida por este producto punto.

b) Muestre que $\langle f,g\rangle_* = f(\vec{v}_g) = g(\vec{v}_f)$.

7. Use integración por partes para mostrar que si u y v tienen derivadas débiles en $L^2\,[a,b]$, entonces, uv tiene derivadas débiles en $L^2\,[a,b]$ y que $(uv)' = u'v + uv'$. Muetre que si p es continuamente diferenciable en $[a,b]$, $p \geq \varepsilon > 0$ y pu es débilmente derivable en $L^2\,[a,b]$, con $u \in L^2\,[a,b]$, entonces u es débilmente derivable en $L^2\,[a,b]$.

8. Sea p continuamente diferenciable en $[a,b]$ con $p \geq \varepsilon > 0$ y sea q continua y no negativa en $[a,b]$.

Pruebe que para toda $f \in L^2\,[a,b]$ existe una única función $u \in H_0^1\,[a,b] \cap H^2\,[a,b]$ tal que

$$\begin{cases} -(pu')' + qu = f \\ \quad u(a) = u(b) = 0 \end{cases}$$

Donde la igualdad se entiende en el sentido de $L^2\,[a,b]$. (Sug.: Use Lax-Milgram y el problema anterior)

9. Considera el operador $l\colon C_0^\infty[a,b] \subset L^2[a,b] \to L^2[a,b]$ dado por

$$lu = -(pu')' + qu$$

con p y q como en el problema 8. Muestre que unp puede exterder el dominio de l de modo que este operador tenga inversa compacta en $L^2\,[a,b]$. (Sug.: Use el problema 8.)

10. Sea l el operador del problema anterior, con p y q como en el problema 8.

a) Muestre que l tiene una base ortonormal de funciones propias en $L^2\,[a,b]$.

b) Use este resultado para mostrar que $\{\operatorname{sen} x, \operatorname{sen} 2x, \operatorname{sen} 3x, \dots\}$ es un conjunto ortogonal completo en $L^2[a,b]$. (Sug.:En a) use el problema 9 para definir l^{-1} y la igualdad $\left\langle l^{-1}f, g\right\rangle = \left\langle l^{-1}f, ll^{-1}g\right\rangle$ para demostrar que l^{-1} es auto-adjunto.)

11. Encunetre la solución del problema mixto

$$\begin{cases} u_t(t,x) = u_{xx}(t,x) \\ u(0,x) = \sin^2 x \\ u(t,0) = u(t,\pi) = 0 \end{cases}$$

(Sug.: Basta calcular el valor del operador $\mathbf{e}^{t\frac{d^2}{dx^2}}$ en la función $\operatorname{sen}^2 x$.)