Solución del problema 2a de tarea 14

El problema. Si $T \in \mathrm{End}(\mathbb{R}^2)$ es ortogonal entonces $\tilde{T}$ (el operador lineal inducido por T en $\mathbb{R}[x,y]$) es ortogonal.

Algunos comentarios preliminares antes de empezar la demostración.

1. Sea $V = \mathbb{R}[x,y]$, el espacio de los polinomios en dos variables x, y. Esto es un espacio vectorial real de dimensión infinita y hemos definido en V (en tarea 13) un producto escalar tal que el conjunto de los monomios $\{x^a y^b | a, b \geq 0\}$ forma una base ortogonal de V con $\|x^a y^b\|^2 = a!b!$. Esta normalización parece misteriosa y requiere algo de motivación (¿porqué no pedir que los monomios sean una base *ortonormal*?).

2. Ofrecemos dos explicaciones. La primera es el problema 3 de la tarea 13. Según este problema, este producto escalar tiene la propiedad que la derivada parcial con respecto a una de las coordenadas es justo el operador adjunto al operador de multiplicación por esta coordenada. Esto es bonito. La segunda justificación a esta definición del producto escalar en V la encontramos en este ejercicio: toda transformación lineal T en $\mathbb{R}^2$ induce una transformación lineal $\tilde{T}$ en V (por cambio de variables en polinomios) y la extraña normalización $\|x^a y^b\|^2 = a!b!$ tiene la propiedad que si T es ortogonal la transformación asociada $\tilde{T}$ en V es ortogonal también.

3. El espacio V es de dimensión infinita, pero es la suma directa ortogonal de subespacios de dimensión finita V_d (polinomios homogeneos de grado d), tal que $\tilde{T}$ deja invariante cada V_d, así que basta demostrar que el operador inducido por $\tilde{T}$ en cada V_d es ortogonal. Es decir, para cada $p, q \in V_d$, basta calcular que $\langle \tilde{T}p, \tilde{T}q \rangle = \langle p, q \rangle$.

4. De hecho, basta demostrar lo anterior cuando p, q son monomios, digamos $p = x^a y^b$ y $q = x^{a'} y^{b'}$, con $a + b = a' + b' = d$. Asi que hay que calcular $\langle \tilde{T}(x^a y^b), \tilde{T}(x^{a'} y^{b'}) \rangle$ y ver que es 0 si $(a,b) \neq (a',b')$, y $a!b!$ cuando $(a,b) = (a',b')$. Al intentar a calcular $\langle \tilde{T}(x^a y^b), \tilde{T}(x^{a'} y^{b'}) \rangle$ se obtiene una expresión bastante complicada, llena de índeces y factoriales. Yo no logré hacer estas cuentas (excepto en el caso $p = q = x^a$), aunque estoy seguro que con suficiente destreza y paciencia se pueden hacer. De todos modos, les ofrezco aquí una demostración indirecta, más sofisticada, involucrando pocas cuentas.

5. La demostración que encontré es un poco larga de explicar y requiere introducir un concepto que no vimos en el curso (formas multilineales). Pero la idea es simple y la explico antes de lanzarnos a la demostración (un resumen muy breve). Identificamos nuestro espacio V_d con otro espacio, S^d, el espacio de las *formas multilineales simétricas de grado d* en $\mathbb{R}^2$. La ventaja de reemplazar V_d por S^d es que S^d es naturalemnte un subespacio de M^d, el espacio de las formas multilineales de grado d (no necesariamente simétricas) en $\mathbb{R}^2$. Ahora M^d, a pesar de ser un espacio mucho más grande que V_d (tiene dimensión 2^d en lugar de $d+1$), tiene un producto escalar más natural por lo que las cuentas son mucho más fáciles que en V_d. La acción de T en $\mathbb{R}^2$, igual que en el caso de V_d, induce una acción "obvia" en M^d, que extiende la acción $\tilde{T}$ en $S^d \subset M^d$, y denotada también por $\tilde{T}$. Es relativamente simple demostrar que la acción de $\tilde{T}$ en M_d es ortogonal. Al restringir el producto escalar en M^d a S^d y usar la identificacón con V_d, obtenemos nuestro producto escalar "extraño" en V_d y así se demuestra que la acción de $\tilde{T}$ en V_d es ortogonal, con respecto al producto escalar que hemos definido en V_d. El punto esencial de

esta demostración es entonces encontrar la correcta identificación $V_d \cong S_d$ que induce nuestro producto escalar en V_d a partir del producto escalar en M^d.

6. En la demostración que sigue el 2 de $\mathbb{R}^2$ no tiene nada especial, y de hecho es más cómodo considerar el caso más general del espacio euclidiano $\mathbb{R}^n$ con coordenadas $x_1, x_2, \dots, x_n$.

La demostración. Sea entonces $E = \mathbb{R}^n$ con su producto escalar estándar. Una forma multilineal de grado d en E, o simplemente *una forma d-lineal*, es una función $f : E^d \to \mathbb{R}$, donde $E^d = E \times \ldots \times E$ (d veces), que es lineal en cada variable separadamente. El espacio de todas las formas d lineales en E se denota por $M^d(E^*)$ o simplemente por M^d. Nota que $M^1(E^*) = E^*$. M^d es un espacio vectorial (bajo la suma y producto por escalar obvios) y contiene como subespacio a las formas multilineales simétricas S^d: son las formas multilineales $f : E^d \to \mathbb{R}$ que no cambian su valor al permutar sus d variables. Más formalmente, denotamos por Σ_d al conjunto de las $d!$ permutaciones de $\{1, \dots, d\}$, y para cada $\sigma \in \Sigma_d$ y $f \in M^d$ definimos $\sigma \cdot f \in M^d$ por $(\sigma \cdot f)(v_1, \dots, v_d) = f(v_{\sigma(1)}, \dots, v_{\sigma(d)})$. Con esta notación, tenemos que $S^d = \{f \in T^d | \sigma \cdot f = f \text{ para todo } \sigma \in \Sigma_d\}$. El producto escalar en E es un ejemplo de un elemento de S^2.

Ejercicio. $(\sigma_1\sigma_2) \cdot f = \sigma_1 \cdot (\sigma_2 \cdot f)$.

Dadas dos formas $f \in M^a$, $g \in M^b$, se define $f \otimes g \in M^{a+b}$ ("producto tensorial") por $(f \otimes g)(v_1, \dots, v_{a+b}) = f(v_1, \dots, v_a)g(v_{a+1}, \dots, v_{a+b})$. Nota que $\otimes$ es un producto bilineal, asociativo, pero no conmutativo. Sea $\alpha_1, \dots, \alpha_n$ una base de E^* y $I = (i_1, \dots, i_d)$, con $1 \leq i_k \leq n$ (multi-índice de grado d). Sea $\alpha_I := \alpha_{i_1} \otimes \cdots \otimes \alpha_{i_d} \in M^d$.

Ejercicio. El conjunto de los α_I es una base de M^d (de aquí vemos que $\dim M^d = n^d$).

Ahora usamos el producto escalar en E para definir un producto escalar en M^d. Tomamos como $\alpha_1, \dots, \alpha_n$ una base dual a una base ortonormal de E (la base canónica por ejemplo), y declaramos al conjunto de los α_I como una base ortonormal de T^d.

Ejercicios.

(1) La definición del producto escalar en M^d no depende de la base ortonormal de E que escogimos (sino solamente del producto escalar en E).

(2) Si $\phi_1, \dots, \phi_d, \psi_1, \dots, \psi_d \in E^*$, entonces

$$\langle \phi_1 \otimes \cdots \otimes \phi_d, \psi_1 \otimes \cdots \otimes \psi_d \rangle = \langle \phi_1, \psi_1 \rangle \langle \phi_2, \psi_2 \rangle \cdots \langle \phi_d, \psi_d \rangle.$$

Ahora para cada $T \in \mathrm{End}(E)$ definimos $\tilde{T} \in \mathrm{End}(M^d)$ por la fórmula $(\tilde{T}f)(v_1, \dots, v_d) := f(Tv_1, \dots, Tv_d)$.

Ejercicios.

(1) $\tilde{T}(\phi_1 \otimes \cdots \otimes \phi_d) = (T^*\phi_1) \otimes \cdots \otimes (T^*\phi_d)$, donde $T^* \in \mathrm{End}(E^*)$ es el operador adjunto, $T^*\phi = \phi \circ T$.

(2) $S^d \subset M^d$ es invariante bajo $\tilde{T}$.

(3) Si $T \in \mathrm{End}(E))$ es ortogonal entonces $\tilde{T} \in \mathrm{End}(M^d)$ es ortogonal.

Ahora sea $V = \mathbb{R}[x_1, \ldots, x_n]$ y $V_d \subset \mathbb{R}[x_1, \ldots, x_n]$ el subespacio de polinomios homogeneos de grado d. Una base para V_d está dada por los monomios de grado d, $\{x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n} | d_1 + \cdots + d_n = d, 0 \leq d_i \leq d\}$. Equipamos a V_d con el producto escalar tal que este conjunto de monomios es ortogonal con $\|x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}\|^2 = d_1! \ldots d_n!$.

La parte principal de la demonstración. Definimos $P : V_d \to M^d$ ("polarización") por

$$P(x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}) = \frac{1}{d!} \sum_{\sigma \in \Sigma_d} \sigma \cdot (x_1^{\otimes d_1} \otimes x_2^{\otimes d_2} \otimes \cdots \otimes x_n^{\otimes d_n}),$$

donde $x_i^{\otimes d_i} := x_i \otimes x_i \otimes \cdots \otimes x_i$ (d_i veces). Aquí estamos considerando a las coordenadas x_i como los elementos de la base dual a la base canónica de E.

Para terminar la demostración del ejercicio tenemos que demostrar lo siguiente:

(A) $P(p) \in S^d$ para todo $p \in V_d$.

(B) P define un ismorfismo lineal $V_d \cong S^d$. La inversa es $f \mapsto f(v, v, \ldots, v)$.

(C) Para todo $T \in \mathrm{End}(E)$, $P \circ \tilde{T} = \tilde{T} \circ P$.

(D) Existe una constante $c > 0$ que depende solamente de d, tal que para todo $p \in V_d$, $\|P(p)\| = c\|p\|$.

Ejercicio. Demostrar incisos (A), (B) y (C).

El único inciso no fácil entre los 4 es el último. Para demostrarlo basta ver que

(1) la imagen de cualquier dos monomios distintos bajo P son elementos ortogonales de M^d;

(2)$\|P(x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n})\|^2 = (d_1! d_2! \cdots d_n!)/d!$ (así que $c = 1/\sqrt{d!}$).

El primer inciso es fácil (ejercicio). Vemos el segundo.

La formula para $P(x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n})$ es un promedio (suma) de $d!$ monomios. Pero muchos se repiten (típicamente). De hecho, vamos a ver que *cada monomio se repite exactamente* $d_1! d_2! \cdots d_n!$ *veces,* por lo que son $d!/d_1! d_2! \cdots d_n!$ monomios distintos. Una vez demostrado esto, podemos calcular

$$\|P(x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n})\|^2 = \left(\frac{1}{d!}\right)^2 \frac{d!}{d_1! d_2! \cdots d_n!} (d_1! d_2! \cdots d_n!)^2 = \frac{d_1! d_2! \cdots d_n!}{d!}.$$

Así que nos queda demostrar que cada uno de los distintos monomio que aparece en la fórmula para $P(x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n})$ se repite $d_1! d_2! \cdots d_n!$ veces. Vamos a examinar primero el monomio que corresponde a $\sigma = id$. Esto es $x_1^{\otimes d_1} \otimes x_2^{\otimes d_2} \otimes \cdots \otimes x_n^{\otimes d_n}$. Es claro que los $\sigma \in \Sigma_d$ que dejan fijo a este monomio son los que dejan invariante el conjunto de los primeros d_1 índices, y el de los siguientes d_2 índices, ...etc. El conjunto $\Sigma' \subset \Sigma_d$ de los elementos que satisface esta propiedad es isomorfo a $\Sigma_{d_1} \times \ldots \times \Sigma_{d_n}$, por lo que tiene $d_1! d_2! \cdots d_n!$ elementos.

De manera similar, si tomamos cualquier elemento de la suma, digamos $\tau \cdot (x_1^{\otimes d_1} \otimes x_2^{\otimes d_2} \otimes \cdots \otimes x_n^{\otimes d_n})$, el conjunto de elementos $\Sigma'' \subset \Sigma_d$ que lo dejan fijo es de la forma $\tau \sigma \tau^{-1}$, donde $\sigma \in \Sigma'$, por lo que Σ' y Σ'' tienen el mismo número de elementos, $d_1! d_2! \cdots d_n!$.

Ejercicio. Demuestra que incisos (A) - (D) implican el ejercicio.

$\square$