Teorema de Perron-Frobenius

Apuntes para Modelos Matemáticos I, Rafael Ortega

1 Enunciado y ejemplos

Suponemos que la matriz $A \in \mathbb{R}^d$ *tiene todos sus coeficientes positivos,*

$$(\star) \quad a_{ij} > 0 \quad \textit{para cada } i, j = 1, \ldots, d.$$

Entonces A *tiene un valor propio* λ_1 *que es positivo y dominante. Además, el vector propio asociado* $w \in \mathbb{R}^d$ *se puede escoger con todas las coordenadas positivas,*

$$Aw = \lambda_1 w, \quad w = \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_d \end{pmatrix},$$

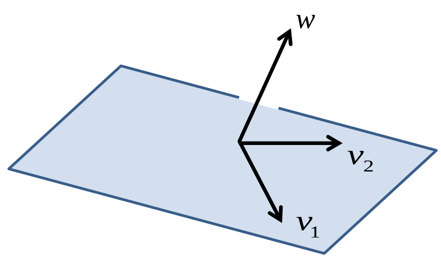
$$\lambda_1 > 0, \quad w_i > 0 \ \textit{ para cada } i = 1, \ldots, d.$$

Ejemplos

1. $A = \begin{pmatrix} 7 & 7 & 7 \\ 7 & 7 & 7 \\ 7 & 7 & 7 \end{pmatrix}$

   En este caso $\lambda_1 = 21$ y podemos elegir $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Comprobamos que λ_1 es dominante ya que solo hay otro valor propio, $\lambda_2 = 0$, que es doble. Los vectores $v_1 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ y $v_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ son linealmente independientes y cumplen $Av_i = 0$.



2. $A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 7 & 7 & 7 \end{pmatrix}$

En este caso $\lambda_1 = 7$ es positivo pero al ser triple no es dominante. No hay vectores propios asociados con todas las coordenadas positivas, el espacio propio asociado es $x + y = 0$. Es un caso en el que el teorema no es aplicable porque la condición $(\star)$ no se cumple.

Una mejora del teorema

El teorema sigue siendo válido cuando $(\star)$ *se cambia por la condición más débil:*

$(\star\star)$ *todos los coeficientes de* A *son no negativos y existe algún entero* $\nu \geq 1$ *para el que la potencia* A^{ν} *tiene todos los coeficientes positivos.*

Ejemplos (continuación)

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

En este caso se cumple ($\star\star$) con $\nu = 2$, pues $A^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Los valores propios de A son $\lambda_1 = \frac{1+\sqrt{5}}{2}$ y $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Como predice el Teorema, λ_1 es positivo y dominante. El vector propio $w = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}$ tiene coordenadas positivas.

Antes de empezar la demostración del teorema vamos a estudiar algunas ideas preliminares sobre el orden vectorial, las aplicaciones lineales y la desigualdad triangular.

2 Preliminares

2.1 El orden vectorial

Definimos

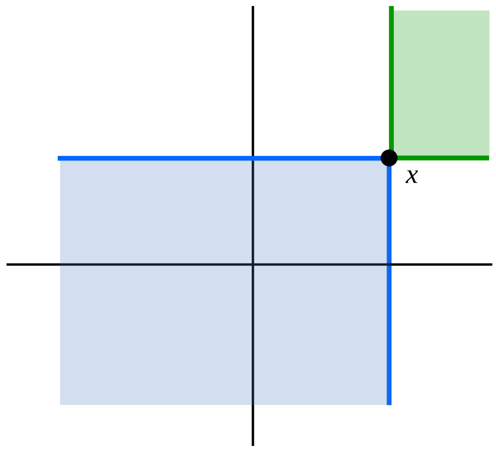
$$\mathbb{R}^d_+ = \{x \in \mathbb{R}^d : \ x = \begin{pmatrix} x_1 \\ \cdots \\ x_d \end{pmatrix}, \ x_i \geq 0, \ i = 1, \ldots, d\}.$$

Para $d = 1$ este conjunto es la semi-recta positiva, para $d = 2$ es el primer cuadrante del plano, para $d = 3$ es el primer octante del espacio,...

Dados $x, y \in \mathbb{R}^d$, la notación $x \geq y$ significa $x - y \in \mathbb{R}^d_+$ o, de manera equivalente,

$$x_i \geq y_i \ \text{ para cada } i = 1, \ldots, d.$$

La notación $x \gg y$ se emplea para la condición más fuerte $x_i > y_i$, $i = 1, \ldots, d$.



Los vectores y en la zona verde cumplen $x \ll y$, mientras que los de la zona azul cumplen $x \gg y$. Los vectores en las zonas sin colorear no se comparan con x. ¿Qué propiedades cumplen los vectores situados en los bordes de las zonas coloreadas?

Ejercicio. *Prueba que si* $x \gg y$ *entonces existe un número* $\epsilon > 0$ *de manera que* $x \geq (1+\epsilon)y$.

Al igual que en el caso de los números, las desigualdades no estrictas van bien con el paso al límite. Dados $x_n, x, \alpha \in \mathbb{R}^d$,

$$\{x_n\}_{n \geq 0} \to x, \quad x_n \geq \alpha \quad \Rightarrow \quad x \geq \alpha.$$

Esto es equivalente a decir que $\mathbb{R}^d_+$ es cerrado en $\mathbb{R}^d$.

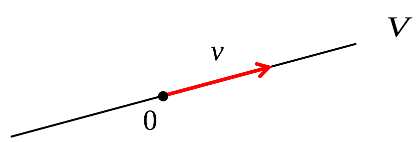
La hipótesis $(\star)$ se puede interpretar como una propiedad de monotonía fuerte de la aplicación lineal descrita por A,

$$x, y \in \mathbb{R}^d, \;\; x \geq y, \; x \neq y \quad \Rightarrow \quad Ax \gg Ay.$$

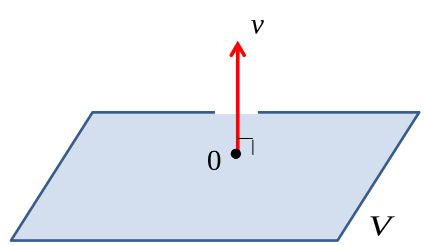
Ejercicio. *Prueba que esta propiedad es equivalente a* $(\star)$.

2.2 La matriz traspuesta y los hiperplanos invariantes

Dada una matriz $A \in \mathbb{R}^{d\times d}$, los vectores propios nos permiten describir las rectas invariantes de la aplicación lineal asociada. Si $Av = \lambda v$, entonces $A(V) \subseteq V$, donde $V = \{cv : \ c \in \mathbb{R}\}$ es la recta vectorial que engendra v.



Por dualidad, los vectores propios reales de la matriz traspuesta A^* nos permiten determinar los hiperplanos invariantes. Si $A^*v = \lambda v$, entonces el hiperplano ortogonal a v es invariante por A.



Lema 1. *Dada* $A \in \mathbb{R}^{d\times d}$, *suponemos que* $v \in \mathbb{R}^d \setminus \{0\}$

cumple $A^*v = \lambda v$ *para algún* $\lambda \in \mathbb{R}$*. Entonces*

$$V = \{x \in \mathbb{R}^d : \ x \perp v\}$$

cumple $A(V) \subseteq V$*.*

Demostración. Usaremos una importante propiedad de la matriz traspuesta:

$$\langle Ax, y\rangle = \langle x, A^*y\rangle, \quad \forall x, y \in \mathbb{R}^d.$$

Dado $x \in V$, sabemos que $\langle x, v\rangle = 0$. Vamos a probar que también Ax es perpendicular a v,

$$\langle Ax, v\rangle = \langle x, A^*v\rangle = \langle x, \lambda v\rangle = 0.$$

2.3 Valores propios simples

Consideramos una aplicación lineal $L : \mathbb{R}^d \to \mathbb{R}^d$. Dada una base $\mathcal{B}$ de $\mathbb{R}^d$, designamos por $M_{\mathcal{B}}$ a la matriz que representa a L en esa base.

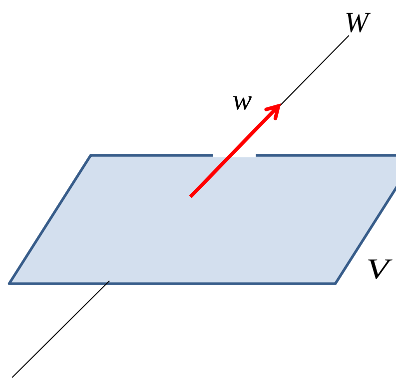
Lema 2. *Suponemos que hay una descomposición*

$$\mathbb{R}^d = V \oplus W$$

donde $W = \mathrm{Ker}(L - \lambda_1 I)$ *es una recta engendrada por un vector propio* $w \in \mathbb{R}^d \setminus \{0\}$ *y* V *es un hiperplano invariante por* L*,* $L(V) \subseteq V$*. Entonces* λ_1 *es un valor propio simple (en sentido algebraico).*

Demostración. Construimos una base $\mathcal{B} = \{w, v_2, \ldots, v_d\}$ de $\mathbb{R}^d$ donde $\{v_2, \ldots, v_d\}$ es una base del sub-espacio V. De $v_i \in V$ se sigue que $Lv_i \in V$, por ser V invariante. Entonces

$$Lv_i = \sum_{j=2}^{d} \alpha_{ji} v_j,$$



con $\alpha_{ij} \in \mathbb{R}^d$. La matriz $M_{\mathcal{B}}$ será

$$M_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \alpha_{22} & \dots & \alpha_{2d} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \alpha_{d2} & \cdots & \alpha_{dd} \end{pmatrix}.$$

La sub-matriz $\hat{M} = \begin{pmatrix} \alpha_{22} & \dots & \alpha_{2d} \\ \cdots & \cdots & \cdots \\ \alpha_{d2} & \cdots & \alpha_{dd} \end{pmatrix}$ representa al operador restringido a V, $L_V : V \to V$, en la base $\hat{\mathcal{B}} = \{v_2, \dots, v_d\}$. Designamos por $p(\lambda)$ al polinomio característico de L y por $\hat{p}(\lambda)$ al de L_V. Entonces, desarrollando por la primera columna de $M_{\mathcal{B}}$,

$$p(\lambda) = \det(M_{\mathcal{B}} - \lambda I) = (\lambda_1 - \lambda)\det(\hat{M} - \lambda I) = (\lambda_1 - \lambda)\hat{p}(\lambda).$$

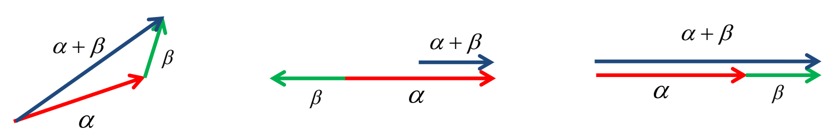
Como λ_1 no es un valor propio de L_V, $\hat{p}(\lambda_1) \neq 0$. Por tanto λ_1 es una raíz simple de $p(\lambda)$.

2.4 Observaciones sobre la desigualdad triangular en $\mathbb{C}$

Dados $\alpha, \beta \in \mathbb{C}$, sabemos que se cumple la desigualdad triangular

$$|\alpha+\beta| \leq |\alpha| + |\beta|.$$

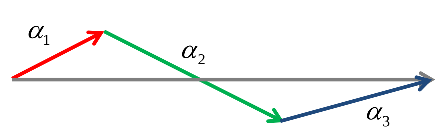
Suponemos $\alpha \neq 0$ y nos preguntamos bajo qué condiciones se da la igualdad. Daremos por cierto lo que sugieren los dibujos; es decir,



$$|\alpha+\beta| = |\alpha| + |\beta| \iff \exists \mu \in \mathbb{R},\ \mu \geq 0 :\ \beta = \mu\alpha.$$

Esta observación se puede extender a cualquier conjunto finito de números complejos.

Lema 3. *Dados $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{C}$ con $\alpha_1 \neq 0$, son equivalentes:*
(i) $|\alpha_1| + |\alpha_2| + \ldots + |\alpha_r| = |\alpha_1 + \alpha_2 + \ldots + \alpha_r|$
(ii) *Para cada $i = 2, \dots, r$ existe $\mu_i \in \mathbb{R},\ \mu_i \geq 0$ tal que $\alpha_i = \mu_i \alpha_1$.*



Demostración. $(ii) \Rightarrow (i)$ Se deja como ejercicio.

$(i) \Rightarrow (ii)$ Sabemos que esta implicación es válida para $r = 2$. Haremos inducción sobre $r \geq 2$. Suponemos que $(i) \Rightarrow (ii)$ es cierta para cualesquiera $r-1$ números y partimos de (i) para r números. Es decir,

$$|\alpha_1| + |\alpha_2| + \ldots + |\alpha_r| = |\alpha_1 + \alpha_2 + \ldots + \alpha_r| = |\alpha_1 + \beta|,$$

donde hemos definido $\beta = \alpha_2 + \ldots \alpha_r$. La desigualdad triangular y esta identidad implican,

$$|\alpha_1 + \beta| \leq |\alpha_1| + |\beta| = |\alpha_1| + |\alpha_2 + \ldots + \alpha_r| \leq$$

$$|\alpha_1| + |\alpha_2| + \ldots + |\alpha_r| = |\alpha_1 + \beta|.$$

Por tanto se cumple $|\alpha_1 + \beta| = |\alpha_1| + |\beta|$. De aquí se sigue la existencia de un número $\mu \geq 0$ tal que $\beta = \mu\alpha_1$. Reescribimos (i) teniendo en cuenta que $1 + \mu$ es positivo,

$$|\alpha_1| + |\alpha_2| + \ldots + |\alpha_r| = |\alpha_1 + \beta| = |\alpha_1 + \mu\alpha_1| = (1 + \mu)|\alpha_1|.$$

Como μ es no negativo deducimos que

$$|\alpha_2| + \ldots + |\alpha_r| = \mu|\alpha_1| = |\mu\alpha_1| = |\beta|.$$

Es decir, se cumple

$$|\alpha_2| + \ldots + |\alpha_r| = |\alpha_2 + \ldots + \alpha_r|$$

y nos encontramos con (i) para $r-1$ números.
No es restrictivo suponer que $\alpha_2 \neq 0$. En otro caso definimos $\mu_2 = 0$ y nos fijamos en α_3.
Por la hipótesis de inducción existen $\gamma_3, \ldots, \gamma_r \geq 0$ tales que $\alpha_j = \gamma_j\alpha_2$, $j \geq 3$. De las identidades anteriores,

$$\mu\alpha_1 = \beta = \alpha_2 + \alpha_3 + \ldots + \alpha_r = (1 + \gamma_3 + \ldots + \gamma_r)\alpha_2.$$

El número $1+\gamma_3+\ldots+\gamma_r$ es positivo y podemos encontrar los coeficientes μ_i,

$$\mu_2 = \frac{\mu}{1+\gamma_3+\ldots+\gamma_r}, \quad \mu_j = \frac{\mu\gamma_j}{1+\gamma_3+\ldots+\gamma_r}, \; j \geq 3.$$

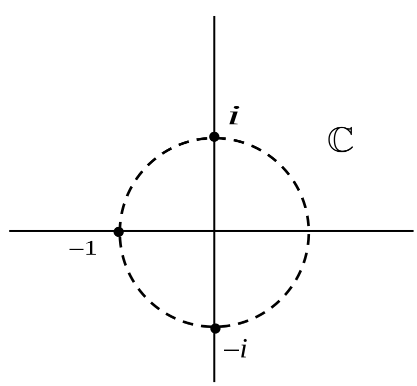
3 Demostración del Teorema

Primero suponemos que se cumple $(\star)$. Al final de la prueba discutiremos los cambios que hay que introducir cuando solo se supone $(\star\star)$.

Comenzamos definiendo la noción de *valor propio de módulo máximo*:

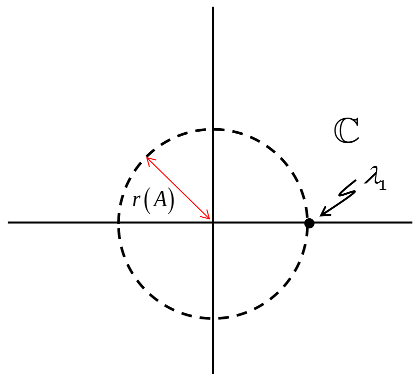
$$\lambda_1 \in \sigma(A), \quad |\lambda_1| = r(A).$$

Algunas matrices tienen varios valores propios con esta propiedad. Por ejemplo, para la matriz $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ todos los valores propios tienen módulo máximo.



El primer paso en la demostración será probar que cuando se cumple ($\star$) hay un único valor propio con módulo máximo. Será consecuencia de la propiedad:

($\clubsuit$) *Si λ_1 es un valor propio de módulo máximo, entonces $\lambda_1 > 0$ y existe $w \in \mathbb{R}^d$, $w \gg 0$ tal que $Aw = \lambda_1 w$.*



Para probar ($\clubsuit$) elegimos $\lambda_1 \in \sigma(A)$ de módulo máximo y un vector propio asociado $v \in \mathbb{C}^d \setminus \{0\}$,

$$Av = \lambda_1 v.$$

En coordenadas,

$$\sum_{j=1}^{d} a_{ij} v_j = \lambda_1 v_i, \quad i = 1, \ldots, d.$$

Después de tomar módulos y usar la desigualdad triangular,

$$|\lambda_1||v_i| = |\sum_{j=1}^{d} a_{ij} v_j| \leq \sum_{j=1}^{d} a_{ij} |v_j|.$$

Aquí hemos usado que los coeficientes de A no son negativos.

Distinguimos dos casos:

(i) todas las desigualdades son estrictas,

$$|\lambda_1||v_i| < \sum_{j=1}^{d} a_{ij}|v_j|, \quad i = 1, \ldots, d$$

(ii) en alguna coordenada se da la igualdad,

$$|\lambda_1||v_i| = \sum_{j=1}^{d} a_{ij}|v_j|, \text{ para algún } i \in \{1, \ldots, d\}.$$

Veamos en primer lugar que el caso (i) es imposible. Definimos el vector

$$V = \begin{pmatrix} |v_1| \\ \cdots \\ |v_d| \end{pmatrix} \in \mathbb{R}_+^d.$$

Si se cumple (i), $AV \gg |\lambda_1|V$. Entonces, por un ejercicio anterior, sabemos que existe $\epsilon > 0$ tal que $AV \geq (|\lambda_1|+\epsilon)V$. La matriz $B = \frac{1}{|\lambda_1|+\epsilon}A$ tiene radio espectral

$$r(B) = \frac{1}{|\lambda_1| + \epsilon} r(A) = \frac{|\lambda_1|}{|\lambda_1| + \epsilon} < 1.$$

Aquí hemos usado que λ_1 tiene módulo máximo. Deducimos que $B^n \to 0$ si $n \to \infty$. Por otra parte sabemos que $BV \geq V$ y, como B preserva el orden, deducimos la cadena de desigualdades

$$\ldots B^{n+1}V \geq B^n V \geq \ldots \geq B^2 V \geq BV \geq V.$$

En particular $B^n V \geq V$ para cada $n \geq 1$. Tomando límites en esta desigualdad y usando la continuidad del producto de matrices, $0 \cdot V = 0 \geq V$. Entonces $V = 0$, lo que es una contradicción con su definición y $v \neq 0$.

Ahora sabemos que la alternativa válida ha de ser (ii) y podemos usar la observación anterior sobre la desigualdad triangular. Para simplificar la notación supondremos que la igualdad se da para la primera coordenada ($i = 1$). Entonces sabemos que se cumple

$$|\lambda_1||v_1| = \sum_{j=1}^{d} a_{1j}|v_j|.$$

Como alguna coordenada v_j es no nula y se cumple $(\star)$,

$$|\lambda_1||v_1| = \sum_{j=1}^{d} a_{1j}|v_j| > 0.$$

Ahora podemos decir que $\lambda_1 \neq 0$ y $v_1 \neq 0$. Usamos el Lema 3 con $\alpha_1 = a_{11}v_1 \neq 0$, $\alpha_j = a_{1j}v_j$, $j \geq 2$. Puesto que se cumple la condición (i) del Lema, existirán nḿeros $\mu_j \geq 0$, $j = 2, \ldots, d$ tales que

$$a_{1j}v_j = \mu_j a_{11} v_1.$$

Hemos obtenido la identidad $v = cw$ con $c = v_1$ y

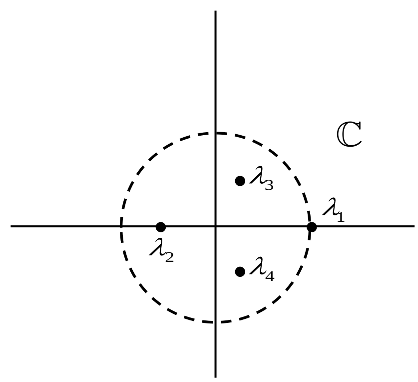
$$V = \begin{pmatrix} 1 \\ \frac{\mu_2 a_{11}}{a_{12}} \\ \cdots \\ \frac{\mu_d a_{11}}{a_{1d}} \end{pmatrix} \in \mathbb{R}_+^d \setminus \{0\}.$$

Como A tiene la propiedad de monotonía fuerte, $Aw \gg 0$. Además, de $w \in \mathrm{Ker}(A - \lambda_1 I)$ se deduce que

$$\lambda_1 w = Aw \gg 0.$$

Como $w_1 = 1$ deducimos de la primera coordenada de esta igualdad que $\lambda_1 > 0$. Finalmente, de $w = \frac{1}{\lambda_1} Aw$ llegamos a $w \gg 0$. Hemos probado que se cumple (♣).

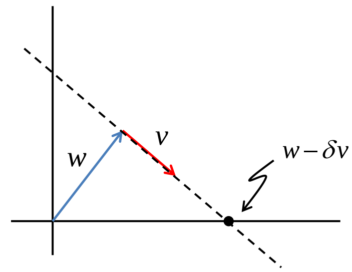
Supongamos ahora que el espectro de la matriz está formado por valores propios distintos $\lambda_1, \lambda_2, \ldots, \lambda_r$ con $r \leq d$ y λ_1 de módulo máximo. Entonces $|\lambda_i| < \lambda_1$ para cada $i = 2, \ldots, r$. Esta propiedad se sigue de (♣) pero todavía no hemos probado que λ_1 sea dominante, falta comprobar que este valor propio es simple en sentido algebraico.



Para ello empezaremos probando que es simple en sentido geométrico; es decir,

$$\dim \operatorname{Ker}(A - \lambda_1 I) = 1.$$

Como w es fuertemente positivo, para cualquier $v \in \mathbb{R}^d \backslash \{0\}$ podemos encontrar $\delta \in \mathbb{R}$ de manera que $w - \delta v \geq 0$ y alguna coordenada de este vector sea nula ($w_i - \delta v_i \geq 0$ para cada i, $w_j - \delta v_j = 0$ para algún j).



Vamos a probar que si $Av = \lambda_1 v$ entonces, en realidad, $w - \delta v = 0$. Esto implica que w y v son linealmente dependientes.

Por reducción al absurdo, si $w - \delta v \geq 0$ con $w - \delta v \neq 0$, como A cumple una propiedad de monotonía fuerte,

$$A(w - \delta v) \gg 0.$$

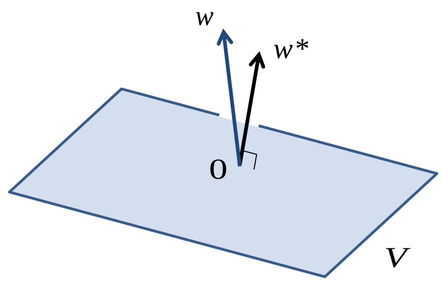
Llegamos a la desigualdad $\lambda_1(w - \delta v) = A(w - \delta v) \gg 0$, que no es compatible con el hecho de que alguna coordenada de $w - \delta v$ sea nula.

Una vez que hemos probado que el espacio propio tiene dimensión 1, nos fijamos en la matriz traspuesta $A^\star$. Esta matriz también cumple la hipótesis ($\star$) y, como consecuencia, la propiedad ($\clubsuit$) es válida para la traspuesta. Las matrices A y $A^\star$ tienen los mismos valores propios, así que λ_1 es el valor propio de módulo máximo para las dos. Encontramos un vector $w^\star \in \mathbb{R}^d$, $w^\star \gg 0$ tal que $A^\star w^\star = \lambda_1 w^\star$.

Por el Lema 1 sabemos que el hiperplano

$$V = \{x \in \mathbb{R}^d : \; x \perp w^\star\}$$

es invariante para A; es decir, $A(V) \subseteq V$.



Vamos a probar que los sub-espacios V y $W = \mathrm{Ker}(A - \lambda_1 I) = \{cw : \ c \in \mathbb{R}\}$ cumplen

$$V \cap W = \{0\}.$$

En otro caso w estaría contenido en V, lo que es absurdo ya que w y $w^\star$ no pueden ser perpendiculares. Al tener ambos vectores todas sus coordenadas positivas, $\langle w, w^\star \rangle > 0$.

Hacemos la suma directa $V \oplus W$, cuya dimensión es $\mathrm{dim}W + \mathrm{dim}V = 1 + d - 1 = d$. Deducimos que $\mathbb{R}^d = V \oplus W$ y aplicamos el Lema 2, así completamos la prueba porque λ_1 es simple también en el sentido algebraico.

4 Modificaciones para la hipótesis $(\star\star)$

En el caso (ii) de la prueba de $(\clubsuit)$ observamos que

$$Av = \lambda_1 v \ \Rightarrow \ A^\nu v = \lambda_1^\nu v.$$

Los valores propios de la matriz A^ν son $\lambda_1^\nu, \lambda_2^\nu, \ldots, \lambda_r^\nu$, de manera que λ_1^ν tiene módulo máximo como valor propio de A^ν. Por otra parte, A^ν cumple $(\star)$ y los argumentos para A se adaptan a esta potencia.

El mismo truco se usa para la prueba de

$$\dim \operatorname{Ker}(A - \lambda_1 I) = 1.$$