Ejercicio 2.3.

Considere en $\mathbb{R}^m$ y $\mathbb{R}^n$ la norma euclidiana. Dada una aplicación lineal $A\colon \mathbb{R}^m \to \mathbb{R}^n$, existe una única aplicación lineal $A^*: \mathbb{R}^n \to \mathbb{R}^m$, llamada la *adjunta* de A, tal que $\langle Ax, y\rangle = \langle x, A^* y\rangle$ para cualesquiera $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Dado $b \in \mathbb{R}^n$, la ecuación $Ax = b$ posee solución $x \in \mathbb{R}^m$ si, y solamente si, b es ortogonal a todo elemento del núcleo de A^*. Concluya que la imagen de A^* y la imagen de A tienen la misma dimensión.

Demostración. Considere en $\mathbb{R}^m$ y $\mathbb{R}^n$ el producto interno canónico. Denotaremos a ambos productos por el mismo símbolo $\langle\cdot,\cdot\rangle$. Considere también en $\mathbb{R}^m$ y $\mathbb{R}^n$ las bases canónicas $\{e_1,\ldots,e_m\}$ y $\{\bar{e}_1,\ldots,\bar{e}_n\}$, respectivamente.

a) Sea $A\colon \mathbb{R}^m \to \mathbb{R}^n$ una transformación lineal y $[a_{ij}]_{n\times m}$ su matriz asociada, respecto a las bases mencionadas. Entonces para todo $j = 1,\ldots,m$, se tienen las igualdades $Ae_j = \sum_{i=1}^n a_{ij}\bar{e}_i$. Sean los vectores $v_1,\ldots,v_n$ definidos como

$$v_i = \sum_{j=1}^{m} a_{ij} e_j\ ,\ i = 1,\ldots,n.$$

Observe que las coordenadas de v_i son las entradas de la i-ésima fila de A. Luego de esto defina la aplicación $A^*: \mathbb{R}^n \to \mathbb{R}^m$ por medio de la regla

$$A^* y = \sum_{i=1}^{n} y_i v_i,$$

para todo $y = (y_1,\ldots.y_n) \in \mathbb{R}^n$. En particular $A^*\bar{e}_i = v_i = \sum_{j=1}^m a_{ij}e_j$. Es claro que A^* es transformación lineal y, además,

$$\langle Ae_j, \bar{e}_k\rangle = \left\langle \sum_{i=1}^{n} a_{ij}\bar{e}_i, \bar{e}_k \right\rangle = a_{kj},$$

mientras que

$$\langle e_j, A^*\bar{e}_k\rangle = \left\langle e_j, \sum_{j=1}^{n} a_{kj} e_j \right\rangle = a_{kj}.$$

Es decir $\langle Ae_j, \bar{e}_k\rangle = \langle e_j, A^*\bar{e}_k\rangle$. Por tanto, para todo $x = \sum_{j=1}^m x_j e_j \in \mathbb{R}^m$ y todo $y = \sum_{k=1}^n y_k \bar{e}_k \in \mathbb{R}^n$ se cumple que

$$\langle Ax, y\rangle = \sum_{j,k} x_j y_k \langle Ae_j, \bar{e}_k\rangle = \sum_{j,k} x_j y_k \langle e_j, A^*\bar{e}_k\rangle = \langle x, A^* y\rangle.$$

Esto muestra que dada una transformación lineal $A\colon \mathbb{R}^m \to \mathbb{R}^n$, existe una transformación lineal $A^*: \mathbb{R}^n \to \mathbb{R}^m$ (la cual es llamada *adjunta de* A), tal que $\langle Ax, y\rangle = \langle x, A^* y\rangle$ para todo $x \in \mathbb{R}^m$ y todo $y \in \mathbb{R}^n$. La aplicación A^* es única, pues si S es otra transformación lineal tal que $\langle Ax, y\rangle = \langle x, Sy\rangle$ para todo $x \in \mathbb{R}^m$ y todo $y \in \mathbb{R}^n$ entonces

$$\langle x, (S - A^*)(\bar{e}_j)\rangle = \langle x, S\bar{e}_j\rangle - \langle x, A^*\bar{e}_j\rangle = \langle Ax, \bar{e}_j\rangle - \langle Ax, \bar{e}_j\rangle = 0.$$

Es decir $(S - A^*)(\bar{e}_j) = 0$. Como esto es válido para cada vector básico $\bar{e}_j$ se concluye fácilmente que $S - A^* = 0$, es decir, $S = A^*$.

b) En lo que sigue, $\ker A^*$ denota al núcleo de A^* mientras que $\operatorname{Img} A$ denota a la imagen de A. Dado un conjunto $X \subseteq \mathbb{R}^k$, el conjunto $X^\perp$ representará al subespacio vectorial formado por todos vectores en $\mathbb{R}^k$ que son ortogonales a todos los elementos de X. Por ejemplo, el tradicional eje Z es un subespacio de $\mathbb{R}^3$ cuyos elementos son ortogonales a todos los vectores contenidos en el plano XY. Luego (eje Z)=(plano XY)$^\perp$.

Considere la ecuación $Ax = b$.

b1) Cuando $b = 0$, la ecuación $Ax = 0$ admite como solución a $x = 0 \in \mathbb{R}^m$. Observe en este caso que $b = 0$ es ortogonal a todo elemento del núcleo de A^*, esto es, $b = 0 \in (\ker A^*)^\perp$.

b2) Ahora suponga $b \neq 0$. Si existe algún $x \in \mathbb{R}^m$ tal que $Ax = b$ entonces, para todo elemento $y \in \ker A^*$ se tiene

$$\left\langle b, y\right\rangle = \left\langle Ax, y\right\rangle = \left\langle x, A^* y\right\rangle = \langle x, 0\rangle = 0.$$

Es decir, si la ecuación $Ax = b$ tiene alguna solución $x \in \mathbb{R}^m$ entonces b es ortogonal a todo elemento del núcleo de A^*, es decir, $b \in (\ker A^*)^\perp$.

Recíprocamente, suponga que $b \in (\ker A^*)^\perp$. En este caso $b \notin \ker A^*$ (de lo contrario b tendría que ser ortogonal a si mismo, lo cual no es posible por ser $b \neq 0$). Observe que

$$z \in \ker A^* \Leftrightarrow \forall\ x \in \mathbb{R}^m \colon\ \langle Ax, z\rangle = \left\langle x, A^* z\right\rangle = 0 \Leftrightarrow z \in (\operatorname{Img} A)^\perp,$$

lo cual muestra que $\ker A^* = (\operatorname{Img} A)^\perp$.

Considere en $\operatorname{Img} A$ una base ortonormal dada por $\{w_1, \ldots, w_k\}$, y considere también el vector $w = \sum_{i=1}^k \langle b, w_i\rangle\, w_i$. Entonces $w \in \operatorname{Img} A \subseteq ((\operatorname{Img} A)^\perp)^\perp$, y como además $b \in (\ker A^*)^\perp = ((\operatorname{Img} A)^\perp)^\perp$ entonces $b - w \in ((\operatorname{Img} A)^\perp)^\perp$. Por otro lado, para todo vector w_j en la base ortonormal considerada de $\operatorname{Img} A$,

$$\left\langle b - w, w_j\right\rangle = \left\langle b - \sum_{i=1}^{k} \langle b, w_i\rangle\, w_i, w_j \right\rangle = \left\langle b, w_j\right\rangle - \left\langle b, w_j\right\rangle = 0,$$

lo cual implica que $b - w \in (\operatorname{Img} A)^\perp$. En conclusión se tiene que

$$b - w \in (\operatorname{Img} A)^\perp \cap ((\operatorname{Img} A)^\perp)^\perp$$

y en particular significa que $b - w$ es ortogonal a si mismo. Como el único vector con esta propiedad es el vector nulo entonces $b - w = 0$, luego $b = w$. Siendo que $w \in \operatorname{Img} A$ entonces $b \in \operatorname{Img} A$. Es decir existe $x \in \mathbb{R}^m$ tal que $Ax = b$.

En conclusión: dado $b \in \mathbb{R}^n$, la ecuación $Ax = b$ admite solución $x \in \mathbb{R}^m$ si, y solo si, b es ortogonal a todo elemento del núcleo de A^*. Dicho de otra manera, b está en la imagen de A si, y solo si, es ortogonal a todo elemento de $\ker A^*$. Esto muestra entonces que $A(\mathbb{R}^m) = (\ker A^*)^\perp$.

c) Del item (b), y usando el teorema del núcleo e imagen, deducimos que

$$\dim(A^*(\mathbb{R}^n)) + \dim(\ker A^*) = n = \dim(\ker A^*)^\perp + \dim(\ker A^*).$$

Por tanto,

$$\dim(A^*(\mathbb{R}^n)) = \dim(\ker A^*)^\perp = \dim(A(\mathbb{R}^m)),$$

es decir, la imagen de A^* y la imagen de A tienen la misma dimensión. ∎

Caso específico

Considere la transformación lineal $A\colon \mathbb{R}^3 \to \mathbb{R}^2$, tal que

$$A(x_1,x_2,x_3) = (x_1+x_2+x_3, -x_1-x_2-x_3).$$

Respecto a las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$, la matriz asociada a A es dada por

$$[a_{ij}]_{2\times 3} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

Usando las filas de esta matriz, definimos los vectores $v_1 = (1,1,1)$ y $v_2 = (-1,-1,-1)$. Entonces definimos $A^*: \mathbb{R}^2 \to \mathbb{R}^3$ por medio de la regla

$$A^*(y_1,y_2) = y_1v_1 + y_2v_2 = (y_1-y_2, y_1-y_2, y_1-y_2).$$

Entonces A^* es la transpuesta de A, pues para todo $x = (x_1,x_2,x_3) \in \mathbb{R}^3$ y todo $y = (y_1,y_2) \in \mathbb{R}^2$ se tiene

$$\begin{aligned}\langle Ax, y\rangle &= \langle (x_1+x_2+x_3, -x_1-x_2-x_3), (y_1,y_2)\rangle = y_1(x_1+x_2+x_3) + y_2(-x_1-x_2-x_3) \\ &= (y_1-y_2)(x_1+x_2+x_3)\end{aligned}$$

mientras que

$$\langle x, A^*y\rangle = \langle (x_1,x_2,x_3), (y_1-y_2, y_1-y_2, y_1-y_2)\rangle = (y_1-y_2)(x_1+x_2+x_3) = \langle Ax, y\rangle.$$

Por otro lado, consideremos la ecuación $Ax = b$, con $b = (r,-r)$, entonces según como se ha definido A se obtiene el sistema

$$\begin{aligned} x_1+x_2+x_3 &= r \\ -x_1-x_2-x_3 &= -r \end{aligned}$$

lo cual nos lleva a la única igualdad $x_1+x_2+x_3 = r$, la misma que tiene infinitas soluciones. Es decir, con $b = (r,-r)$ el sistema $Ax = b$ tiene (infinitas) solución.
De la definición de A^*, su núcleo es dado por todos aquellos vectores $y = (y_1,y_2) \in \mathbb{R}^2$ tales que $A^*y = 0$, es decir, todos aquellos vectores $y \in \mathbb{R}^2$ tales que $(y_1-y_2, y_1-y_2, y_1-y_2) = (0,0,0)$, luego

$$\ker(A^*) = \{t(1,1)\colon\ t \in \mathbb{R}\}.$$

y se observa que para todo $t \in \mathbb{R}$,

$$\langle b, t(1,1)\rangle = \langle (r,-r), (t,t)\rangle = 0,$$

es decir, b es ortogonal al núcleo de A^*. Si b no es ortogonal a $\ker A^*$, por ejemplo $b = (1,2)$ entonces el sistema $Ax = b$ no tiene solución, pues en este caso tendríamos las igualdades

$$\begin{aligned} x_1+x_2+x_3 &= 1 \\ -x_1-x_2-x_3 &= 2 \end{aligned}$$

lo que nos llevaría al absurdo $1 = -2$. En conclusión $Ax = b$ tiene solución si, y solo si, b es ortogonal a $\ker A^*$, tal como se afirma en el Ejercicio 2.3.