Scaling along a line in the plane

One way to compute the matrix for a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$ is to determine the effect of the linear transformation on the standard basis vectors of $\mathbf{R}^n$. Sometimes, however, it may be useful or even easier to compute the matrix for a linear transformation from the matrices for other linear transformations. That is the approach we will take in these exercises.

The linear transformation of interest here (scaling along a line in the plane) will be denoted by f and its matrix will be labeled A. So, $f : \mathbf{R}^2 \to \mathbf{R}^2$ and $f(\vec{x}) = A\,\vec{x}$ for any $\vec{x} \in \mathbf{R}^2$. Let L be the line through the origin in $\mathbf{R}^2$ that is parallel to the unit vector $\hat{u}$ which is at the angle θ with respect to the horizontal. f is a scaling along L by the factor s. The real numbers θ and s are arbitrary but fixed. f depends on both parameters. We will compute A by two different methods. As you might expect, θ and s will appear in A when we compute it.

(1) We have already examined the case of h, a horizontal scaling by the factor s and easily found that its matrix is $H = \begin{bmatrix} s & 0 \\ 0 & 1 \end{bmatrix}$. One would hope that A and H are simply related and we could obtain A from H. In fact, if we didn't have the coordinate axes for reference, we would be unable to distinguish between the effects of the transformations f and h. The two are related to one another through rotation by θ. Let r be the counter-clockwise rotation by θ. We know its matrix, too. It is $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. The relationship between f and h is this: $f = r \circ h \circ r^{-1}$. In other words, $f(\vec{x}) = r(h(r^{-1}(\vec{x})))$ for any $\vec{x} \in \mathbf{R}^2$. This asserts that we can achieve the transformation f in three steps. First, apply r^{-1}, then apply h, and finally apply r.

a. Explain why it follows that $A = R\,H\,R^{-1}$.

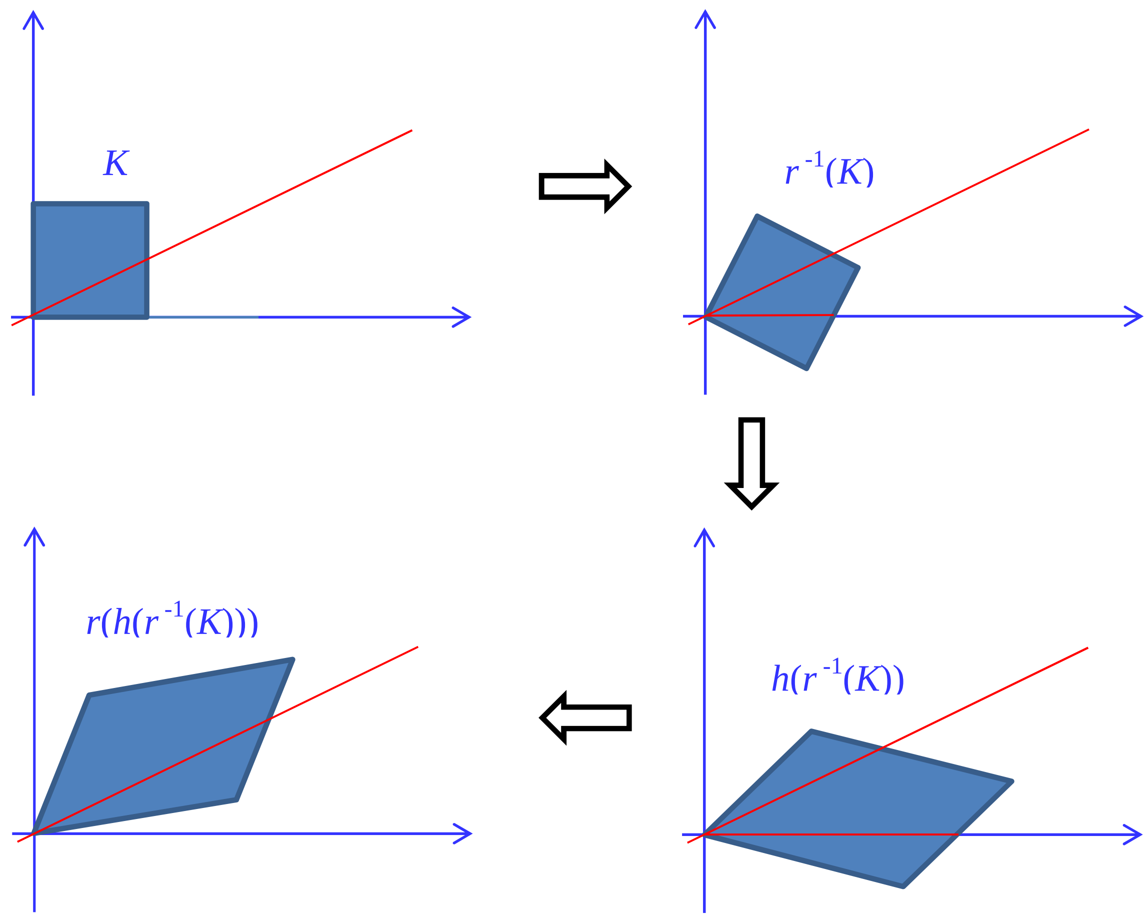
The matrix for a composite transformation is the product of the matrices corresponding to the constituent transformations applied (reading from right to left) in the same order.

b. What is R^{-1}? [There is an easy and a hard way to determine R^{-1}.]

The inverse of rotation through the angle θ is rotation through the angle $-\theta$. The is easily verified by computing the product of the two matrices. So, we have

$$R^{-1} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

c. Geometrically convince the reader why $f = r \circ h \circ r^{-1}$ is true by a sequence of careful drawings. Let K be the unit square, $K = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid 0 \le x \le 1, 0 \le y \le 1 \right\}$. Make four drawings of K, $r^{-1}(K)$, $h(r^{-1}(K))$ and $r(h(r^{-1}(K)))$ in succession. For these illustrations, use $s = 2$ and let L be the line with equation $x = 2y$. $g(K)$ denotes the set obtained by applying g to all the vectors in K.



d. Compute and simplify A.

$$A = R\,H\,R^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} s & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} s\cdot\cos^2\theta + \sin^2\theta & (s-1)\sin\theta\cos\theta \\ (s-1)\sin\theta\cos\theta & \cos^2\theta + s\cdot\sin^2\theta \end{bmatrix}.$$

(2) In the second method, we resolve the action of f into its effect on that part of a vector that is parallel to L and the action of f on that part that is orthogonal to L. Let p be the linear transformation that is projection onto L and let P be

its matrix. Let $\vec{x}$ be <u>any</u> vector in the plane. $p(\vec{x})$ is the projection of $\vec{x}$ along L and it is given by $p(\vec{x}) = (\hat{u}\cdot\vec{x})\hat{u}$ as shown in class.

e. The vector $\vec{x} - (\hat{u}\cdot\vec{x})\hat{u}$ is the projection of $\vec{x}$ orthogonal to L. It is that part of $\vec{x}$ that remains after removing its projection onto L.

There is no question or problem here.

f. Show $\vec{x} - (\hat{u}\cdot\vec{x})\hat{u}$ is orthogonal to $\hat{u}$. Also, show that the matrix for projection orthogonal to L is $I - P$.

Two vectors are orthogonal if and only if their dot (inner or scalar) product is zero.

$\hat{u}\cdot\left(\vec{x} - (\hat{u}\cdot\vec{x})\hat{u}\right) = \hat{u}\cdot\vec{x} - \hat{u}\cdot\left((\hat{u}\cdot\vec{x})\hat{u}\right) = \hat{u}\cdot\vec{x} - (\hat{u}\cdot\vec{x})(\hat{u}\cdot\hat{u}) = \hat{u}\cdot\vec{x} - \hat{u}\cdot\vec{x} = 0$ shows $\vec{x} - (\hat{u}\cdot\vec{x})\hat{u}$ is orthogonal to $\hat{u}$. Now, the image of an arbitrary vector $\vec{x}$ under projection orthogonal to L is $\vec{x} - (\hat{u}\cdot\vec{x})\hat{u} = I\vec{x} - P\vec{x} = (I-P)\vec{x}$. So, $I - P$ is the matrix for projection orthogonal to L.

Clearly, $I = P + (I-P)$ and $\vec{x} = I\vec{x} = P\vec{x} + (I-P)\vec{x}$. This identity allows us to resolve any vector into its projections parallel to L and orthogonal to L.

g. What does multiplication by A do to a vector parallel to L? What does multiplication by A do to a vector orthogonal to L? Use this to simplify the identity $A\vec{x} = AP\vec{x} + A(I-P)\vec{x}$.

$$A\vec{x} = AP\vec{x} + A(I-P)\vec{x} = sP\vec{x} + (I-P)\vec{x} = \left(I + (s-1)P\right)\vec{x}$$

h. If B and C are two 2×2 matrices and $B\vec{x} = C\vec{x}$ f or all vectors $\vec{x}$ in the plane, then $B = C$. Prove this.

Substituting the first standard basis vector and then the second standard basis vector for $\vec{x}$ leads us to conclude that the first and second columns of B and C are the same.

i. Use parts g and h to derive a formula for A in terms of I and P.

$A = I + (s-1)P$.

j. First determine $\hat{u}$ and then compute P in terms of θ.

$\hat{u} = \begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}$. The first and second columns of P are the images, respectively, of the standard basis vectors $\hat{e}_1$ and $\hat{e}_2$. So,

$$P = \left[(\hat{u}\cdot\hat{e}_1)\hat{u} \mid (\hat{u}\cdot\hat{e}_2)\hat{u}\right] = \left[(\cos\theta)\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix} \middle| (\sin\theta)\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}\right] = \begin{bmatrix}\cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta\end{bmatrix}$$

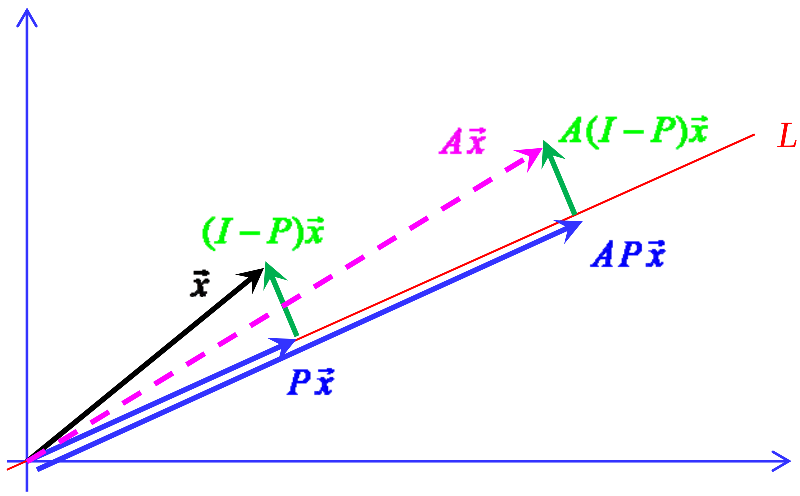
k. Compute A from parts i and j. Simplify.

$$A = I + (s-1)P = \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} + (s-1)\begin{bmatrix}\cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta\end{bmatrix}$$

$$= \begin{bmatrix}1+(s-1)\cos^2\theta & (s-1)\cos\theta\sin\theta \\ (s-1)\cos\theta\sin\theta & 1+(s-1)\sin^2\theta\end{bmatrix} = \begin{bmatrix}s\cos^2\theta + \sin^2\theta & (s-1)\cos\theta\sin\theta \\ (s-1)\cos\theta\sin\theta & \cos^2\theta + s\sin^2\theta\end{bmatrix}.$$

Scaling along a line in the plane

l. Illustrate the construction in part i as follows. Draw and label the following in $\mathbf{R}^2$: the coordinate axes, the line L with equation $x = 2y$, the vector $\vec{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, the unit vector $\hat{u}$ parallel to L, the vectors $P\vec{x}$ and $(I-P)\vec{x}$. Suppose that $s = 2$. Now, label and draw $AP\vec{x}$ and $A(I-P)\vec{x}$. Finally, label and construct $A\vec{x}$ from the previous two vectors.



m. Compare the two expressions you have for A from parts d and k.

The matrices are identical.

For each of the following, simplify A and discuss your result.

n. $s = 1$

A scaling by a factor of 1 is no change at all. Indeed, if we substitute $s = 1$, we have $A = I$.

o. $s = -1$

This should be a reflection across a line orthogonal to L. In fact, if $s = -1$, we obtain $A = I - 2P = \begin{bmatrix} -\cos^2\theta + \sin^2\theta & -2\cos\theta\sin\theta \\ -2\cos\theta\sin\theta & \cos^2\theta - \sin^2\theta \end{bmatrix} = -\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$.

This is a reflection across a line at the angle θ relative to the horizontal followed by a reflection across the origin. This is the same as a reflection across $L^\perp$, a line through the origin orthogonal to L.

p. $\theta = 0$

Not surprisingly, substitution leads us to $A = H = \begin{bmatrix} s & 0 \\ 0 & 1 \end{bmatrix}$.

q. $\theta = \pi/2$

Scaling along a line in the plane

We obtain $A = \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}$ which is a vertical scaling by the factor s.

r. $\theta = \pi$

In this case, we find $A = \begin{bmatrix} s & 0 \\ 0 & 1 \end{bmatrix} = H$, a horizontal scaling by s, as before. This makes sense since rotating L by π yields L.

s. $s = 0$ [Hint: in part i, you computed P, the matrix for projection onto L. Now compute and simplify the matrix $I - P$, the projection orthogonal to L.]

Substituting $s = 0$, we find $A = \begin{bmatrix} \sin^2\theta & -\cos\theta\sin\theta \\ -\cos\theta\sin\theta & \cos^2\theta \end{bmatrix}$ and this is projection onto the unit vector $\hat{u}^{\perp} = \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$ that is orthogonal to $\hat{u}$. In other words, this is projection onto $L^{\perp}$.