Linear Algebra

Linear Transformations and their Matrices

Source: Introduction to linear algebra by Gilbert Strang

When a matrix A multiplies a vector $\boldsymbol{v}$, it "transforms" $\boldsymbol{v}$ into another vector $A\boldsymbol{v}$ ***In goes*** $\boldsymbol{v}$, ***out comes*** $A\boldsymbol{v}$. This transformation follows the same idea as a function. In goes a number x, out comes $f(x)$. For one vector $\boldsymbol{v}$ or one number x, we multiply by the matrix or we evaluate the function. The deeper goal is to see all $\boldsymbol{v}$'s at once. We are transforming the whole space when we multiply every $\boldsymbol{v}$ by A.

Start again with a matrix A. It transforms $\boldsymbol{v}$ to $A\boldsymbol{v}$. It transforms $\boldsymbol{w}$ to $A\boldsymbol{w}$. Then we *know* what happens to $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$. There is no doubt about $A\boldsymbol{u}$, it has to equal $A\boldsymbol{v} + A\boldsymbol{w}$. Matrix multiplication $T(\boldsymbol{v}) = A\boldsymbol{v}$ gives a ***linear transformation***:

DEFINITION A transformation T assigns an output $T(\boldsymbol{v})$ to each input vector $\boldsymbol{v}$. The transformation is ***linear*** if it meets these requirements for all $\boldsymbol{v}$ and $\boldsymbol{w}$:

(a) $T(\boldsymbol{v} + \boldsymbol{w}) = T(\boldsymbol{v}) + T(\boldsymbol{w})$ (b) $T(c\boldsymbol{v}) = cT(\boldsymbol{v})$ for all c.

If the input is $\boldsymbol{v} = \mathbf{0}$, the output must be $T(\boldsymbol{v}) = \mathbf{0}$. We combine (a) and (b) into one:

$$\textbf{\textit{Linearity:}} \quad T(c\boldsymbol{v} + d\boldsymbol{w}) \quad \textbf{\textit{must equal}} \quad cT(\boldsymbol{v}) + dT(\boldsymbol{w}).$$

Again I test matrix multiplication: $A(c\boldsymbol{v} + d\boldsymbol{w}) = cA\boldsymbol{v} + dA\boldsymbol{w}$ is *true*.

A linear transformation is highly restricted. Suppose T adds $\boldsymbol{u}_0$ to every vector. Then $T(\boldsymbol{v}) = \boldsymbol{v} + \boldsymbol{u}_0$ and $T(\boldsymbol{w}) = \boldsymbol{w} + \boldsymbol{u}_0$. This isn't good, or at least *it isn't linear*. Applying T to $\boldsymbol{v} + \boldsymbol{w}$ produces $\boldsymbol{v} + \boldsymbol{w} + \boldsymbol{u}_0$. That is not the same as $T(\boldsymbol{v}) + T(\boldsymbol{w})$:

$$\boldsymbol{v} + \boldsymbol{w} + \boldsymbol{u}_0 \qquad \text{is different from} \qquad T(\boldsymbol{v}) + T(\boldsymbol{w}) = \boldsymbol{v} + \boldsymbol{u}_0 + \boldsymbol{w} + \boldsymbol{u}_0.$$

The exception is when $\boldsymbol{u}_0 = \mathbf{0}$. The transformation reduces to $T(\boldsymbol{v}) = \boldsymbol{v}$. This is the ***identity transformation*** (nothing moves, as in multiplication by I). That is certainly linear. In this case the input space $\mathbf{V}$ is the same as the output space $\mathbf{W}$.

Example 1 Choose a fixed vector $\boldsymbol{a} = (1, 3, 4)$, and let $T(\boldsymbol{v})$ be the dot product $\boldsymbol{a} \cdot \boldsymbol{v}$:

The input is $\boldsymbol{v} = (v_1, v_2, v_3)$. The output is $T(\boldsymbol{v}) = \boldsymbol{a} \cdot \boldsymbol{v} = v_1 + 3v_2 + 4v_3$.

This is linear. The inputs $\boldsymbol{v}$ come from three-dimensional space, so $\mathbf{V} = \mathbf{R}^3$. The outputs are just numbers, so the output space is $\mathbf{W} = \mathbf{R}^1$. We are multiplying by the row matrix $A = [1 \ \ 3 \ \ 4]$. Then $T(\boldsymbol{v}) = A\boldsymbol{v}$.

You will get good at recognizing which transformations are linear. If the output involves squares or products or lengths, v_1^2 or $v_1 v_2$ or $\|\boldsymbol{v}\|$, then T is not linear.

Example 2 The length $T(\boldsymbol{v}) = \|\boldsymbol{v}\|$ is not linear. Requirement (a) for linearity would be $\|\boldsymbol{v} + \boldsymbol{w}\| = \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$. Requirement (b) would be $\|c\boldsymbol{v}\| = c\|\boldsymbol{v}\|$. Both are false!

Not (a): The sides of a triangle satisfy an *inequality* $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$.

Not (b): The length $\| -\boldsymbol{v}\|$ is not $-\|\boldsymbol{v}\|$. For negative c, we fail.

Example 3 (Important) T is the transformation that *rotates every vector by* $30°$. The domain is the xy plane (where the input vector $\boldsymbol{v}$ is). The range is also the xy plane (where the rotated vector $T(\boldsymbol{v})$ is). We described T without mentioning a matrix: just rotate the plane by $30°$.

Is rotation linear? *Yes it is.* We can rotate two vectors and add the results. The sum of rotations $T(\boldsymbol{v}) + T(\boldsymbol{w})$ is the same as the rotation $T(\boldsymbol{v} + \boldsymbol{w})$ of the sum. The whole plane is turning together, in this linear transformation.

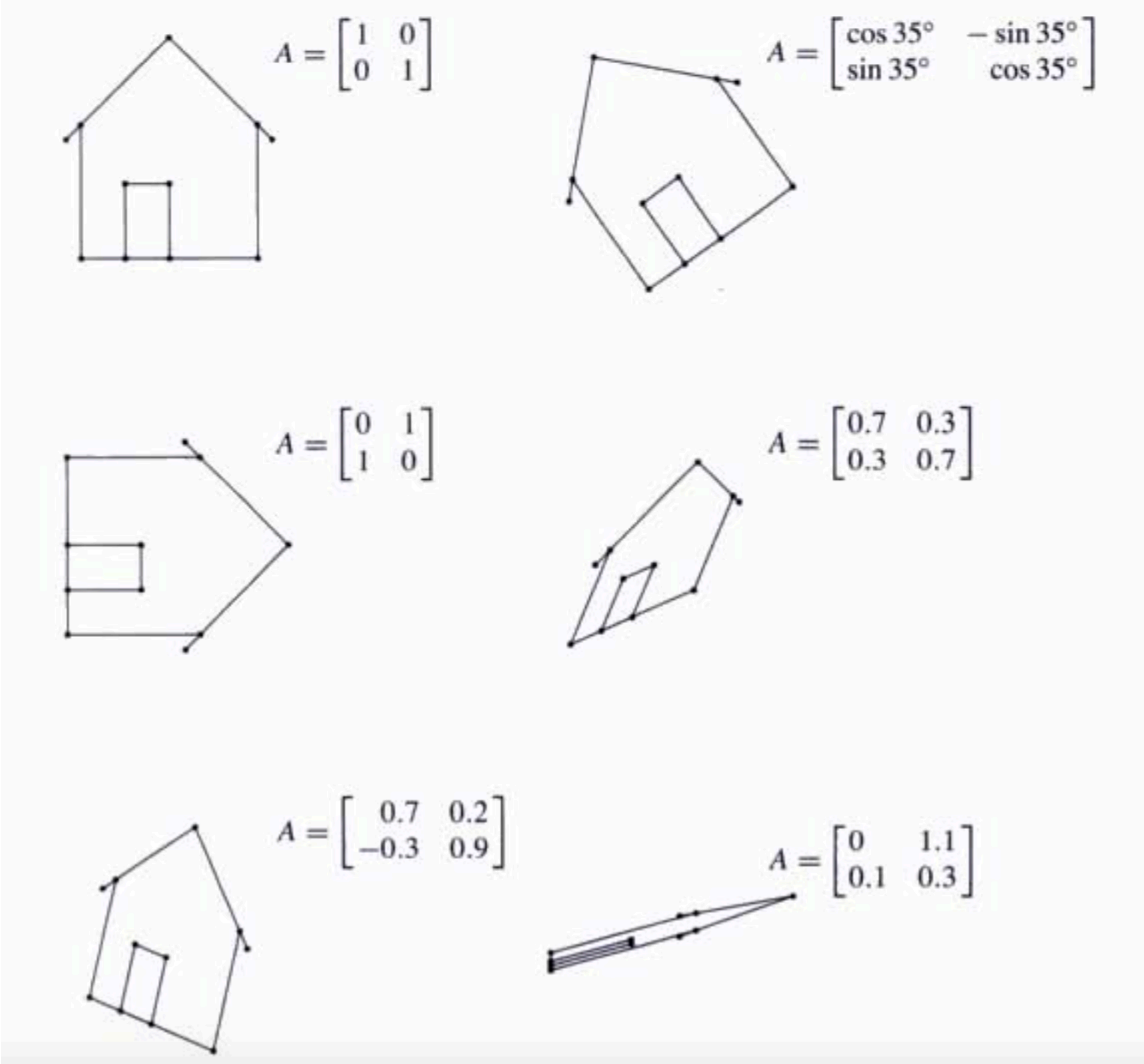
Note Transformations have a language of their own. Where there is no matrix, we can't talk about a column space. But the idea can be rescued and used. The column space consisted of all outputs $A\boldsymbol{v}$. The nullspace consisted of all inputs for which $A\boldsymbol{v} = \mathbf{0}$. Translate those into "range" and "kernel":

Range of T = set of all outputs $T(\boldsymbol{v})$: corresponds to column space

Kernel of T = set of all inputs for which $T(\boldsymbol{v}) = \mathbf{0}$: corresponds to nullspace.

Example 6 Suppose A is an *invertible matrix*. The kernel of T is the zero vector; the range $\mathbf{W}$ equals the domain $\mathbf{V}$. Another linear transformation is multiplication by A^{-1}. This is the ***inverse transformation*** T^{-1}, which brings every vector $T(\boldsymbol{v})$ back to $\boldsymbol{v}$:

$$T^{-1}(T(\boldsymbol{v})) = \boldsymbol{v} \quad \text{matches the matrix multiplication} \quad A^{-1}(A\boldsymbol{v}) = \boldsymbol{v}.$$


$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$A = \begin{bmatrix} \cos 35^\circ & -\sin 35^\circ \\ \sin 35^\circ & \cos 35^\circ \end{bmatrix}$
$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
$A = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$
$A = \begin{bmatrix} 0.7 & 0.2 \\ -0.3 & 0.9 \end{bmatrix}$
$A = \begin{bmatrix} 0 & 1.1 \\ 0.1 & 0.3 \end{bmatrix}$

The next pages assign a matrix to every linear transformation. For ordinary column vectors, the input $\boldsymbol{v}$ is in $\mathbf{V} = \mathbf{R}^n$ and the output $T(\boldsymbol{v})$ is in $\mathbf{W} = \mathbf{R}^m$. The matrix for this transformation T will be m by n.

The standard basis vectors for $\mathbf{R}^n$ and $\mathbf{R}^m$ lead to a standard matrix for T. Then $T(\boldsymbol{v}) = A\boldsymbol{v}$ in the normal way. But these spaces also have other bases, so the same T is represented by other matrices. A main theme of linear algebra is to choose the bases that give the best matrix.

Key idea of this section

When we know $T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)$ for the basis vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, linearity produces $T(\boldsymbol{v})$ for every other vector $\boldsymbol{v}$.

Reason Every input $\boldsymbol{v}$ is a unique combination $c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$ of the basis vectors. Since T is a linear transformation (here is the moment for linearity), the output $T(\boldsymbol{v})$ must be the same combination of the known outputs $T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)$:

$$\begin{array}{c} \textit{Suppose } v = c_1 v_1 + \cdots + c_n v_n. \\ \textit{Then linearity requires } T(v) = c_1 T(v_1) + \cdots + c_n T(v_n). \end{array} \tag{1}$$

Example 1 Suppose T transforms $\boldsymbol{v}_1 = (1, 0)$ to $T(\boldsymbol{v}_1) = (2, 3, 4)$. Suppose the second basis vector $\boldsymbol{v}_2 = (0, 1)$ goes to $T(\boldsymbol{v}_2) = (5, 5, 5)$. If T is linear from $\mathbf{R}^2$ to $\mathbf{R}^3$ then its "standard matrix" is 3 by 2. Those outputs go into its columns:

$$A = \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \end{bmatrix}. \qquad T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) \quad \text{is} \quad \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}.$$

Construction of the Matrix

Now we construct a matrix for any linear transformation. Suppose T transforms the space $\mathbf{V}$ (n-dimensional) to the space $\mathbf{W}$ (m-dimensional). We choose a basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ for $\mathbf{V}$ and a basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ for $\mathbf{W}$. The matrix A will be m by n. To find its first column, apply T to the first basis vector $\boldsymbol{v}_1$:

$$\textit{\textbf{T}}(\boldsymbol{v}_1) \textit{\textbf{ is a combination}} \quad a_{11}\boldsymbol{w}_1 + \cdots + a_{m1}\boldsymbol{w}_m \quad \textit{\textbf{of the output basis for W.}}$$

These numbers $a_{11}, \ldots, a_{m1}$ go into the first column of A. Transforming $\boldsymbol{v}_1$ to $T(\boldsymbol{v}_1)$ matches multiplying $(1, 0, \ldots, 0)$ by A. It yields that first column of the matrix.

7A Each linear transformation T from $\mathbf{V}$ to $\mathbf{W}$ is represented by a matrix A (after the bases are chosen for $\mathbf{V}$ and $\mathbf{W}$). The jth column of A is found by applying T to the jth basis vector $\boldsymbol{v}_j$:

$$T(\boldsymbol{v}_j) = \text{combination of basis vectors of } \mathbf{W} = a_{1j}\boldsymbol{w}_1 + \cdots + a_{mj}\boldsymbol{w}_m. \tag{5}$$

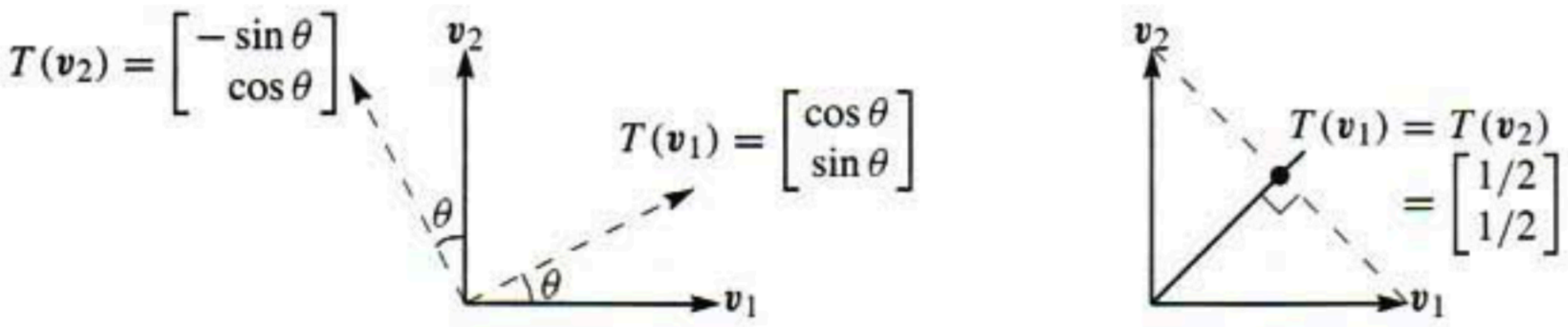


Figure 7.2 Rotation by θ and projection onto the 45° line.

Example 5 T rotates every plane vector by the same angle θ. Here $\mathbf{V} = \mathbf{W} = \mathbf{R}^2$. Find the rotation matrix A. The answer depends on the basis!

Solution The standard basis is $\boldsymbol{v}_1 = (1, 0)$ and $\boldsymbol{v}_2 = (0, 1)$. To find A, apply T to those basis vectors. In Figure 7.2a, they are rotated by θ. *The first vector* $(1, 0)$ ***swings around to*** $(\cos\theta, \sin\theta)$. This equals $\cos\theta$ times $(1, 0)$ plus $\sin\theta$ times $(0, 1)$. Therefore those numbers $\cos\theta$ and $\sin\theta$ go into the *first column* of A:

$$\begin{bmatrix} \cos\theta & \\ \sin\theta & \end{bmatrix} \text{ shows column 1} \quad A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ shows both columns.}$$

For the second column, transform the second vector $(0, 1)$. The figure shows it rotated to $(-\sin\theta, \cos\theta)$. *Those numbers go into the second column.* Multiplying A times $(0, 1)$ produces that column, so A agrees with T.

Example 6 (*Projection*) Suppose T projects every plane vector onto the 45° line. Find its matrix for two different choices of the basis. We will find two matrices.

Solution Start with a specially chosen basis, not drawn in Figure 7.2. The basis vector $\boldsymbol{v}_1$ is along the 45° line. *It projects to itself.* From $T(\boldsymbol{v}_1) = \boldsymbol{v}_1$, the first column of A contains 1 and 0. The second basis vector $\boldsymbol{v}_2$ is along the perpendicular line (135°). *This basis vector projects to zero.* So the second column of A contains 0 and 0:

$$\textbf{Projection } A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{when } \mathbf{V} \text{ and } \mathbf{W} \text{ have the } 45^\circ \text{ and } 135^\circ \text{ basis.}$$

Now take the standard basis $(1, 0)$ and $(0, 1)$. Figure 7.2b shows how $(1, 0)$ projects to $(\frac{1}{2}, \frac{1}{2})$. That gives the first column of A. The other basis vector $(0, 1)$ also projects to $(\frac{1}{2}, \frac{1}{2})$. So the standard matrix for this projection is A:

$$\textbf{Projection } A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \text{for the } \textit{same} \; T \text{ and the standard basis.}$$

Problem

Which of these transformations is not linear? The input is $\boldsymbol{v} = (v_1, v_2)$:

(a) $T(\boldsymbol{v}) = (v_2, v_1)$ (b) $T(\boldsymbol{v}) = (v_1, v_1)$ (c) $T(\boldsymbol{v}) = (0, v_1)$

(d) $T(\boldsymbol{v}) = (0, 1)$.