Linear Transformations

CS6015 / LARP

ACK : Linear Algebra and Its Applications - Gilbert Strang

Introduction

We know how a matrix moves subspaces around when we multiply by A. The nullspace goes into the zero vector. All vectors go into the column space, since Ax is always a combination of the columns. You will soon see something beautiful—that A takes its row space into its column space, and on those spaces of dimension r it is 100 percent invertible. That is the real action of A. It is partly hidden by nullspaces and left nullspaces, which lie at right angles and go their own way (toward zero).

What matters now is what happens inside the space—which means inside n −dimensional space, if A is $n\ by\ n$. That demands a closer look.

Suppose x is an n −dimensional vector. When A multiplies x, it transforms that vector into a new vector Ax. This happens at every point x of the n −dimensional space $\boldsymbol{R}^n$. The whole space is transformed, or "*mapped into itself*," by the matrix A.

$$A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$$

1. A multiple of the identity matrix, $A = cI$, ***stretches*** every vector by the same factor c. The whole space expands or contracts (or somehow goes through the origin and out the opposite side, when c is negative).

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

2. A ***rotation*** matrix turns the whole space around the origin. This example turns all vectors through 90°, transforming every point (x, y) to $(-y, x)$.

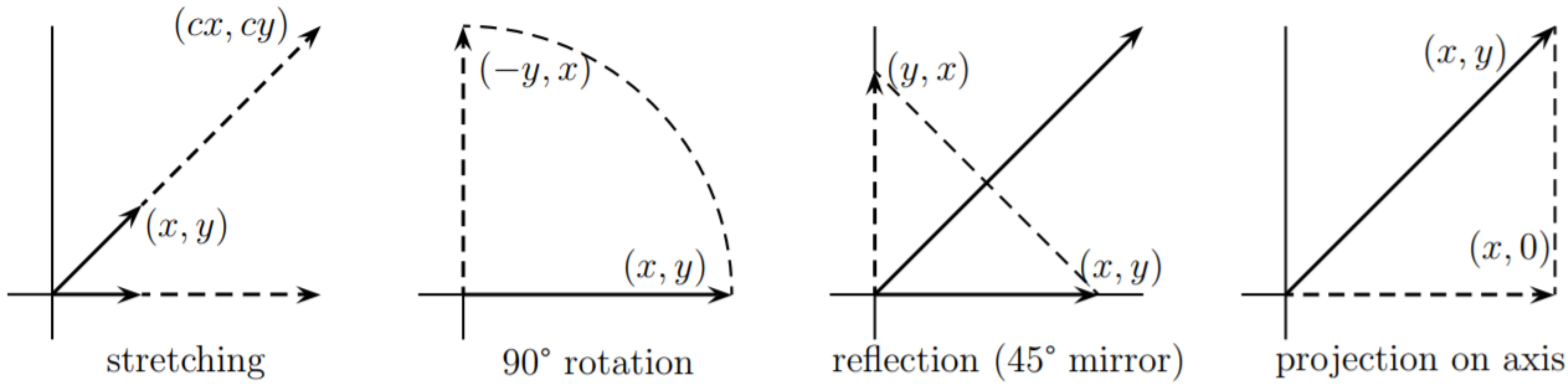


Figure 2.9: Transformations of the plane by four matrices.

Those examples could be lifted into three dimensions. There are matrices to stretch the earth or spin it or reflect it across the plane of the equator (forth pole transforming to south pole). There is a matrix that projects everything onto that plane (both poles to the center).

It is also important to recognize that matrices cannot do everything, and some transformations $T(x)$ are not possible with Ax:

i. It is impossible to move the origin, since $A0 = 0$ for every matrix.

ii. If the vector x goes to x' , then $2x$ must go to $2x'$. in general cx must go to cx' , since $A(cx) = c(Ax)$.

iii. If the vectors x and y go to x' and y' , then their sum $x + y$ must go to $x' + y'$—since $A(x + y) = Ax + Ay$.

Matrix multiplication imposes those rules on the transformation. The second rule contains the first (take $c = 0$ to get $A0 = 0$). We saw rule iii. in action when (4,0) was reflected across the 45° line. It was split into (2,2) + (2,−2) and the two parts were reflected separately. The same could be done for projections: split, project separately, and add the projections. These rules apply to **any transformation that comes from a matrix.**

Their importance has earned them a name: Transformations that obey rules i.–iii. are called **linear transformations**. The rules can be combined into one requirement:

2T For all numbers c and d and all vectors x and y, matrix multiplication satisfies the rule of linearity:

$$A(cx + dy) = c(Ax) + d(Ay). \tag{1}$$

Every transformation $T(x)$ that meets this requirement is a ***linear transformation***.

Any matrix leads immediately to a linear transformation. The more interesting question is in the opposite direction: *Does every linear transformation lead to a matrix?*

The object of this section is to find the answer, yes. This is the foundation of an approach to linear algebra—starting with property (1) and developing its consequences—that is much more abstract than the main approach in this book. We preferred to begin directly with matrices, and now we see how they represent linear transformations.

A transformation need not go from $\boldsymbol{R}^n$ to the same space $\boldsymbol{R}^n$. It is absolutely permitted to transform vectors in $\boldsymbol{R}^n$ to vectors in a different space $\boldsymbol{R}^m$. That is exactly what is done by an $m\ by\ n$ matrix! The original vector x has n components, and the transformed vector Ax has m components. The rule of linearity is equally satisfied by rectangular matrices, so they also produce linear transformations.

Having gone that far, there is no reason to stop. The operations in the linearity condition (1) are addition and scalar multiplication, but x and y need not be column vectors in $\boldsymbol{R}^n$. Those are not the only spaces. By definition, any vector space allows the combinations $cx + dy$—the "vectors" are x and y, but they may actually be polynomials or matrices or functions $x(t)$ and $y(t)$. As long as the transformation satisfies equation (1), it is linear.

We take as examples the spaces $\boldsymbol{P}_n$, in which the vectors are polynomials $p(t)$ of degree n. They look like $p = a_0 + a_1 t + \cdots + a_n t^n$, and the dimension of the vector space is $n + 1$ (because with the constant term, there are $n + 1$ coefficients).

Example 1. The operation of ***differentiation***, $A = d/dt$, is linear:

$$Ap(t) = \frac{d}{dt}(a_0 + a_1 t + \cdots + a_n t^n) = a_1 + \cdots + n a_n t^{n-1}. \tag{2}$$

The nullspace of this A is the one-dimensional space of constants: $da_0/dt = 0$. The column space is the n-dimensional space $\mathbf{P}_{n-1}$; the right-hand side of equation (2) is always in that space. The sum of nullity $(= 1)$ and rank $(= n)$ is the dimension of the original space $\mathbf{P}_n$.

Example 2. ***Integration*** from 0 to t is also linear (*it takes* $\boldsymbol{P}_n$ *to* $\boldsymbol{P}_{n+1}$):

$$Ap(t) = \int_0^t (a_0 + \cdots + a_n t^n)\,dt = a_0 t + \cdots + \frac{a_n}{n+1} t^{n+1}. \tag{3}$$

This time there is no nullspace (except for the zero vector, as always!) but integration does not produce all polynomials in $\mathbf{P}_{n+1}$. The right side of equation (3) has no constant term. Probably the constant polynomials will be the left nullspace.

Example 3. ***Multiplication*** by a fixed polynomial like $2+3t$ is linear:

$$Ap(t) = (2+3t)(a_0 + \cdots + a_n t^n) = 2a_0 + \cdots + 3a_n t^{n+1}.$$

Again this transforms $\mathbf{P}_n$ to $\mathbf{P}_{n+1}$, with no nullspace except $p = 0$.

In these examples (and in almost all examples), linearity is not difficult to verify. It hardly even seems interesting. If it is there, it is practically impossible to miss. Nevertheless, it is the most important property a transformation can have[1]. Of course most transformations are not linear—for example, to square the polynomial ($Ap = p^2$), or to add 1 ($Ap = p+1$), or to keep the positive coefficients ($A(t-t^2) = t$). It will be linear transformations, and *only those*, that lead us back to matrices.

---

[1]Invertibility is perhaps in second place as an important property.

Transformations Represented by Matrices

Linearity has a crucial consequence: **If we know Ax for each vector in a basis, then we know Ax for each vector in the entire space.** Suppose the basis consists of the n vectors $x_1, \ldots, x_n$. Every other vector x is a combination of those particular vectors (they span the space). Then linearity determines Ax:

$$\textbf{Linearity} \qquad \text{If} \quad x = c_1x_1 + \cdots + c_nx_n \quad \text{then} \quad Ax = c_1(Ax_1) + \cdots + c_n(Ax_n). \qquad (4)$$

The transformation $T(x) = Ax$ has no freedom left, after it has decided what to do with the basis vectors. The rest is determined by linearity. The requirement (1) for two vectors x and y leads to condition (4) for n vectors $x_1, \ldots, x_n$. The transformation does have a free hand with the vectors in the basis (they are independent). When those are settled, the transformation of every vector is settled.

Example 4. What linear transformation takes x_1 and x_2 to Ax_1 and Ax_2?

$$x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{goes to} \quad Ax_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}; \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{goes to} \quad Ax_2 = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}.$$

It must be multiplication $T(x) = Ax$ by the matrix

$$A = \begin{bmatrix} 2 & 4 \\ 3 & 6 \\ 4 & 8 \end{bmatrix}.$$

Starting with a different basis (1,1) and (2,−1), this same A is also the only linear transformation with

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \\ 12 \end{bmatrix} \quad \text{and} \quad A\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Next we find matrices that represent differentiation and integration. **First we must decide on a basis**. For the polynomials of degree 3 there is a natural choice for the four basis vectors:

$$\textbf{Basis for } \mathbf{P}_3 \qquad p_1 = 1, \quad p_2 = t, \quad p_3 = t^2, \quad p_4 = t^3.$$

That basis is not unique (it never is), but some choice is necessary and this is the most convenient. The derivatives of those four basis vectors are $0, 1, 2t, 3t^2$:

$$\textbf{Action of } d/dt \qquad Ap_1 = 0, \quad Ap_2 = p_1, \quad Ap_3 = 2p_2, \quad Ap_4 = 3p_3. \tag{5}$$

"d/dt" is acting exactly like a matrix, but which matrix? Suppose we were in the usual four-dimensional space with the usual basis—the coordinate vectors p_1=(1,0,0,0), p_2=(0,1,0,0), p_3=(0,0,1,0), p_4=(0,0,0,1). The matrix is decided by equation (5):

$$\textbf{Differentiation matrix} \qquad A_{\text{diff}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Differentiation matrix $$A_{\text{diff}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ap_1 is its first column, which is zero. Ap_2 is the second column, which is p_1. Ap_3 is $2p_2$ and Ap_4 is $3p_3$. The nullspace contains p_1 (the derivative of a constant is zero). The column space contains p_1, p_2, p_3 (the derivative of a cubic is a quadratic). The derivative of a combination like $p = 2 + t - t^2 - t^3$ is decided by linearity, and there is nothing new about that—it is the way we all differentiate. It would be crazy to memorize the derivative of every polynomial.

The matrix can differentiate that $p(t)$, because matrices build in linearity!

$$\frac{dp}{dt} = Ap \longrightarrow \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -3 \\ 0 \end{bmatrix} \longrightarrow 1 - 2t - 3t^2.$$

In short, *the matrix carries all the essential information*. If the basis is known, and the matrix is known, then the transformation of every vector is known.

The coding of the information is simple. To transform a space to itself, one basis is enough. A transformation from one space to another requires a basis for each.

2U Suppose the vectors $x_1, \ldots, x_n$ are a basis for the space $\mathbf{V}$, and vectors $y_1, \ldots, y_m$ are a basis for $\mathbf{W}$. Each linear transformation T from $\mathbf{V}$ to $\mathbf{W}$ is represented by a matrix A. The jth column is found by applying T to the jth basis vector x_j, and writing $T(x_j)$ as a combination of the y's:

$$\textbf{Column } j \textbf{ of } A \qquad T(x_j) = Ax_j = a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m. \tag{6}$$

For the differentiation matrix, column 1 came from the first basis vector $p1 = 1$. Its derivative is zero, so column 1 is zero. The last column came from $\left(\frac{d}{dt}\right) t^3 = 3t^2$. Since $3t^2 = 0p_1 + 0p_2 + 3p_3 + 0p_4$, the last column contained 0, 0, 3. 0. The rule (6) constructs the matrix, a column at a time.

We do the same for integration. That goes from cubics to quartics, transforming $\mathbf{V} = \mathbf{P}_3$ into $\mathbf{W} = \mathbf{P}_4$, so we need a basis for $\mathbf{W}$. The natural choice is $y_1 = 1$, $y_2 = t$, $y_3 = t^2$, $y_4 = t^3$, $y_5 = t^4$, spanning the polynomials of degree 4. The matrix A will be m by n, or 5 by 4. It comes from applying integration to each basis vector of $\mathbf{V}$:

$$\int_0^t 1\,dt = t \quad \text{or} \quad Ax_1 = y_2, \quad \ldots, \quad \int_0^t t^3\,dt = \frac{1}{4}t^4 \quad \text{or} \quad Ax_4 = \frac{1}{4}y_5.$$

$$\textbf{Integration matrix} \qquad A_{\text{int}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}.$$

Differentiation and integration are *inverse operations*. Or at least integration *followed* by differentiation brings back the original function. To make that happen for matrices, we need the differentiation matrix from quartics down to cubics, which is 4 by 5:

$$A_{\text{diff}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix} \qquad \text{and} \qquad A_{\text{diff}} A_{\text{int}} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}.$$

Differentiation is a *left-inverse* of integration. Rectangular matrices cannot have two-sided inverses! In the opposite order, $A_{int}A_{diff} = I$ cannot be true. The $5\ by\ 5$ product has zeros in column 1. The derivative of a constant is zero. In the other columns $A_{int}A_{diff}$ is the identity and the integral of the derivative of t^n is t^n .

Rotations Q, Projections P, and Reflections H

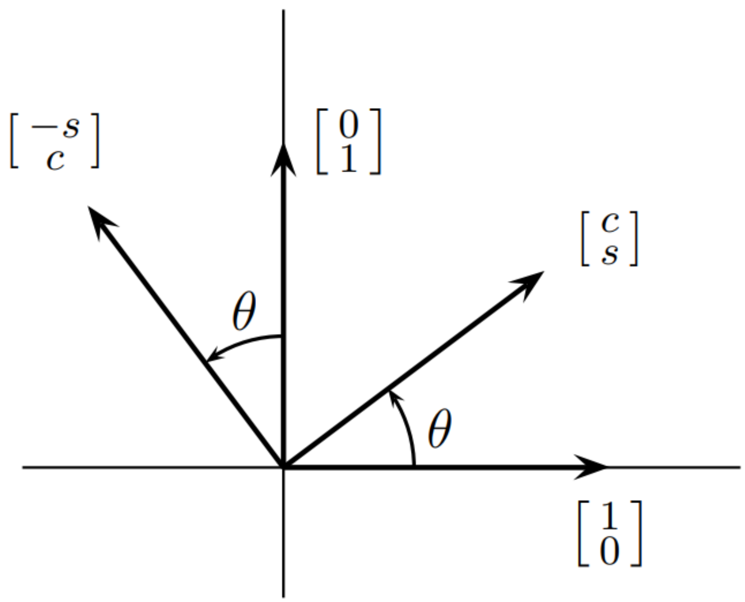
This section began with 90° rotations, projections onto the x-axis, and reflections through the 45° line. Their matrices were especially simple:

$$Q = \underset{\text{(rotation)}}{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}} \qquad P = \underset{\text{(projection)}}{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}} \qquad H = \underset{\text{(reflection)}}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}.$$

The underlying linear transformations of the $x - y$ plane are also simple. But rotations through other angles, projections onto other lines, and reflections in other mirrors are almost as easy to visualize, They are still linear transformations, provided that the origin is fixed: $A0 = 0$. They *must* be represented by matrices. Using the natural basis $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we want to discover those matrices.

1. Rotation

Figure 2.10 shows rotation through an angle θ. It also shows the effect on the two basis vectors. The first one goes to $(\cos\theta, \sin\theta)$, whose length is still 1; it lies on the "θ-line." The second basis vector (0,1) rotates into $(-\sin\theta, \cos\theta)$. By rule (6), those numbers go into the columns of the matrix (we use c and s for $\cos\theta$ and $\sin\theta$). This family of rotations Q_θ is a perfect chance to test the correspondence between transformations and matrices:



$$R = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

$$P = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$$

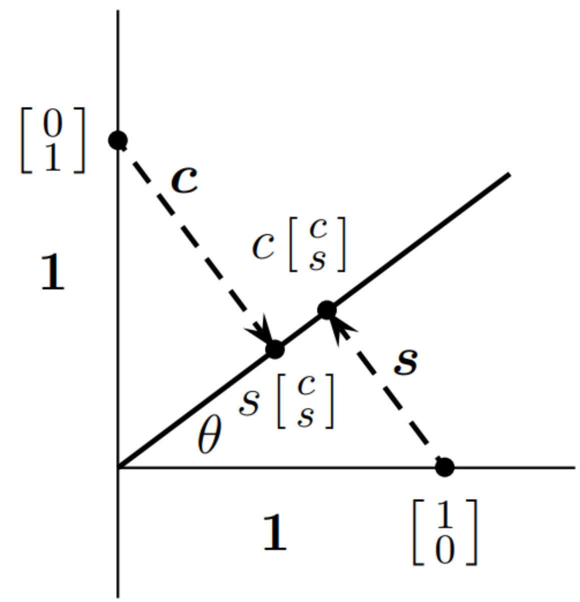


Figure 2.10: Rotation through θ (left). Projection onto the $\boldsymbol{\theta}$-line (right).

*Does the **inverse** of Q_θ equal $Q_{-\theta}$* (rotation backward through θ)? *Yes.*

$$Q_\theta Q_{-\theta} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

*Does the **square** of Q_θ equal $Q_{2\theta}$* (rotation through a double angle)? *Yes.*

$$Q_\theta^2 = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} c^2 - s^2 & -2cs \\ 2cs & c^2 - s^2 \end{bmatrix} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}.$$

*Does the **product** of Q_θ and Q_φ equal $Q_{\theta+\varphi}$* (rotation through θ then φ)? *Yes.*

$$Q_\theta Q_\varphi = \begin{bmatrix} \cos\theta\cos\varphi - \sin\theta\sin\varphi & \cdots \\ \sin\theta\cos\varphi + \cos\theta\sin\varphi & \cdots \end{bmatrix} = \begin{bmatrix} \cos(\theta+\varphi) & \cdots \\ \sin(\theta+\varphi) & \cdots \end{bmatrix}.$$

The last case contains the first two. The inverse appears when φ is $-\theta$, and the square appears when φ is $+\theta$. All three questions were decided by trigonometric identities (and they give a new way to remember those identities). It was no accident that all the answers were yes. *Matrix multiplication is defined exactly so that **the product of the matrices corresponds to the product of the transformations***.

2V Suppose A and B are linear transformations from $\boldsymbol{V}$ to $\boldsymbol{W}$ and from $\boldsymbol{U}$ to $\boldsymbol{V}$. Their product AB starts with a vector u in $\boldsymbol{U}$, goes to Bu in $\boldsymbol{V}$, and finishes with ABu in $\boldsymbol{W}$. This "*composition*" AB is again a linear transformation (from $\boldsymbol{U}$ to $\boldsymbol{W}$). Its matrix is the product of the individual matrices representing A and B.

For $A_{diff}A_{int}$, the composite transformation was the identity (and $A_{int}A_{diff}$ annihilated all constants). For rotations, the order of multiplication does not matter. Then $\boldsymbol{U} = \boldsymbol{V} = \boldsymbol{W}$ is the $x - y$ plane, and $Q_\theta Q_\phi$ is the same as $Q_\phi Q_\theta$. For a rotation and a reflection, the order makes a difference.

Technical note: To construct the matrices, we need bases for $\boldsymbol{V}$ and $\boldsymbol{W}$, and then for $\boldsymbol{U}$ and $\boldsymbol{V}$. By keeping the same basis for $\boldsymbol{V}$, the product matrix goes correctly from the basis in $\boldsymbol{U}$ to the basis in $\boldsymbol{W}$. If we distinguish the transformation A from its matrix (call that $[A]$), then the product rule $2\boldsymbol{V}$ becomes extremely concise: $[AB] = [A][B]$. The rule for multiplying matrices was totally determined by this requirement—it must match the product of linear transformations.

2. Projection

Figure 2.10 also shows the projection of (1,0) onto the θ-line. The length of the projection is $c = \cos\theta$. Notice that the point of projection is not (c, s), as I mistakenly thought; that vector has length 1 (it is the rotation), so we must multiply by c. Similarly the projection of (0,1) has length s, and falls at $s(c, s) = (cs, s^2)$, that gives the second column of the projection matrix P:

$$\textbf{Projection onto } \boldsymbol{\theta}\textbf{-line} \qquad P = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}.$$

This matrix has no inverse, because the transformation has no inverse. Points on the perpendicular line are projected onto the origin; that line is the nullspace of P. Points on the θ-line are projected to themselves! Projecting twice is the same as projecting once, and $P^2 = P$:

$$P^2 = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}^2 = \begin{bmatrix} c^2(c^2+s^2) & cs(c^2+s^2) \\ cs(c^2+s^2) & s^2(c^2+s^2) \end{bmatrix} = P.$$

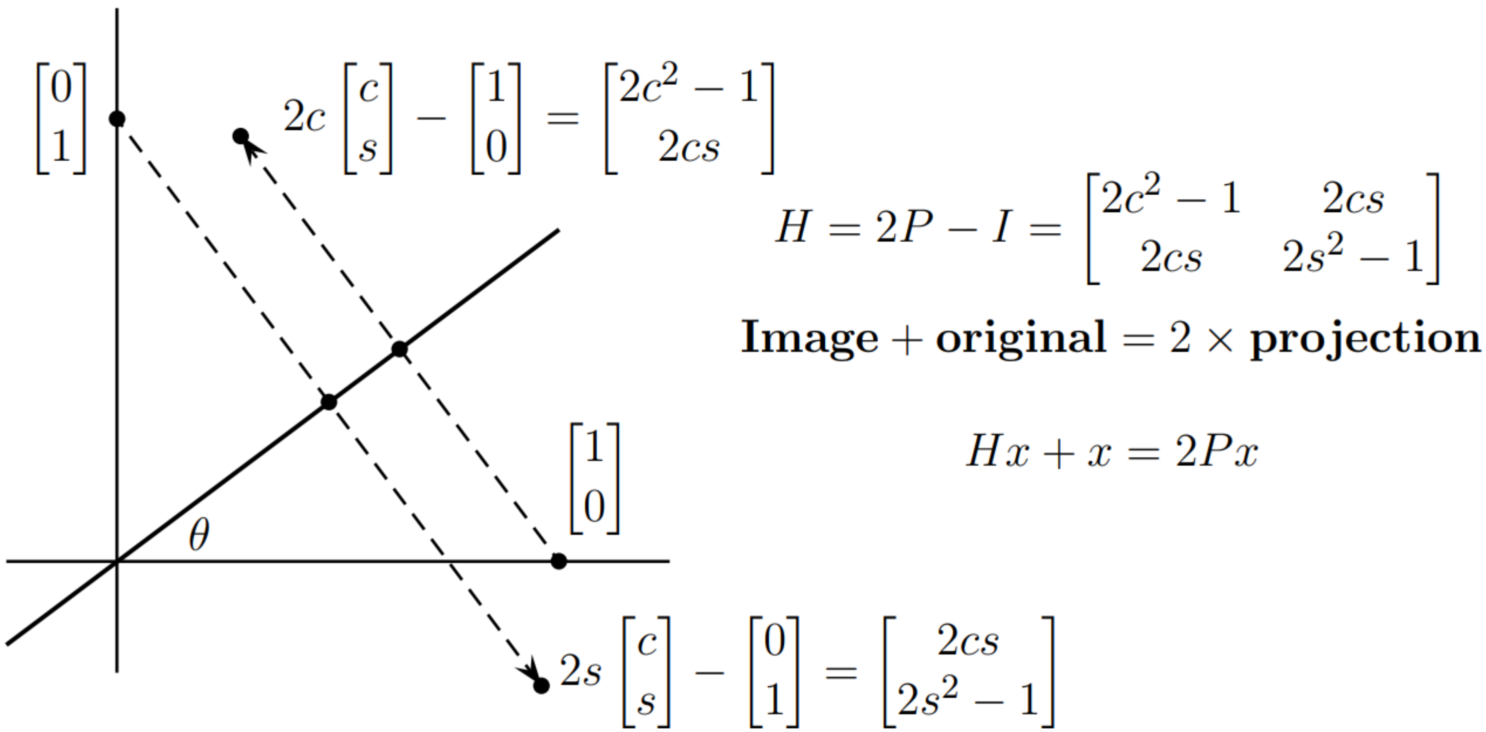


Figure 2.11: Reflection through the θ-line: the geometry and the matrix.

Of course $c^2+s^2=\cos^2\theta+\sin^2\theta=1$. ***A projection matrix equals its own square***.

3. Reflection

Figure 2.11 shows the reflection of (1,0) in the θ-line. The length of the reflection equals the length of the original, as it did after rotation—but here the θ-line stays where it is. The perpendicular line reverses direction; all points go straight through the mirror, Linearity decides the rest.

$$\textbf{Reflection matrix} \qquad H = \begin{bmatrix} 2c^2-1 & 2cs \\ 2cs & 2s^2-1 \end{bmatrix}.$$

This matrix H has the remarkable property $H^2 = I$. **Two reflections bring back the original**. A reflection is its own inverse, $H = H^{-1}$, which is clear from the geometry but less clear from the matrix. One approach is through the relationship of reflections to projections: $H = 2P - I$. This means that $Hx + x = 2Px$—the image plus the original equals twice the projection. It also confirms that $H^2 = I$:

$$H^2 = (2P-I)^2 = 4P^2 - 4P + I = I, \qquad \text{since} \quad P^2 = P.$$

Other transformations Ax can increase the length of x; stretching and shearing are in the exercises. Each example has a matrix to represent it—which is the main point of this section. But there is also the question of choosing a basis, and we emphasize that the matrix depends on the choice of basis. Suppose the first basis vector is **on the θ-line** and the second basis vector is **perpendicular**:

i. The projection matrix is back to $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. This matrix is constructed as always: its first column comes from the first basis vector (projected to itself). The second column comes from the basis vector that is projected to zero.

ii. For reflections, that same basis gives $H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. The second basis vector is reflected onto its negative, to produce this second column. The matrix H is still $2P - I$ when the same basis is used for H and P.

iii. For rotations, the matrix is not changed. Those lines are still rotated through θ, and $Q = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$ as before.

The whole question of choosing the best basis is absolutely central. The goal is to make the matrix diagonal, as achieved for P and H. To make Q diagonal requires complex vectors, since all real vectors are rotated.

We mention here the effect on the matrix of a change of basis, while the linear transformation stays the same. **The matrix** A (or Q or P or H) **is altered to** $S^{-1}AS$. Thus a single transformation is represented by different matrices (via different bases, accounted for by S). The theory of eigenvectors will lead to this formula $S^{-1}AS$, and to the best basis.