Geometry of Elementary Operations

Notation:

We write $\mathbf{x} = \mathbf{B}x$ for a vector in a geometrical space, perhaps the plane or our familiar 3-dimensional space, wherein basis $\mathbf{B}$ connects $\mathbf{x}$ with the column vector x of its components via an invertible linear transformation. Similarly $\mathbf{w}^T = w^T\mathbf{B}^{-1}$ is a linear functional acting upon vectors in the geometrical space and represented by a row vector w^T via the same basis $\mathbf{B}$. Hence, $\mathbf{w}^T\mathbf{x} = w^Tx$; in fact, all expressions involving geometrical entities like $\mathbf{x}$ and $\mathbf{w}^T$ in this note are computed by replacing those entities' symbols by their numerical representatives x and w^T. So long as we stick with just one basis $\mathbf{B}$ we might as well avoid the bother of typing boldface characters; whenever we see " x " we shall let the context determine whether this refers to the column vector x or to the geometrical vector $\mathbf{x}$.

The equation " r^Tx = constant " confines x to a *line* if the whole space is 2-dimensional, a *plane* if 3-dimensional, an *hyperplane* if the dimension exceeds 3 . In this note the word " plane " will be used as an abbreviation for " line or plane or hyper-plane."

Elementary Projectors:

For any vector c and functional r^T such that $r^Tc \neq 0$, let $P := cr^T/r^Tc$. Then P *projects* any vector z onto $y = Pz$ as follows: Find the plane, in the family of parallel planes whose equations are " r^Tx = constant ", that passes through z ; its equation is " $r^Tx = r^Tz$ ". Find the line through the origin o parallel to c ; its parametric equation is " $x = c\lambda$ " wherein λ runs through all real scalars. The line and plane intersect at $y = c(r^Tz/r^Tc) = Pz$ because y is parallel to c and $r^Ty = r^Tz$. Some writers call P a "projector", some a "projection".

In short, elementary projector P collapses the whole vector space to a line through o parallel to c , and the direction of collapse is parallel to the planes " r^Tx = constant ". *Draw a picture*.

Note that $I \neq P = P^2 \neq O$. This relation characterizes every projector. Another projector is $Q := I - P = I - cr^T/r^Tc$. *Confirm that it satisfies* $I \neq Q = Q^2 \neq O$ *too*. It collapses the whole space into the plane " $r^Tx = 0$ ", and the direction of collapse is parallel to c , first because $r^TQz = o^Tz = 0$ and second because $z - Qz = Pz$ is parallel to c . *Draw a picture again*.

Unlike the projectors in most introductory texts, ours need not be orthogonal; we can use any direction c not in the plane " $r^Tx = 0$ ", and project parallel to the plane onto c via P , or else parallel to c onto the plane via Q . Our *Affine* vector space need not be *Euclidean*.

The projectors P and Q are called *complementary* (not spelt " compl<u>i</u>mentary ") because $P+Q = I$. They decompose an arbitrary vector z into two components; Pz is the component parallel to c , and Qz is the component in the plane " $r^Tx = 0$ ".

Exercise: Verify $PQ = QP = O$; what does this mean geometrically?

Elementary Reflectors:

A *reflector* is a linear operator W that satisfies $W^2 = I \neq W$. Some writers call W a "reflection" or, if British, a "reflexion". An example is $-I$, which reflects through the origin o (it reflects z to $-z$). All *elementary reflectors* have the form $W := I - 2cr^T/r^Tc$ with $r^Tc \neq 0$, so $W = I - 2P = 2Q - I$. Every elementary reflector leaves a plane unchanged; the plane's equation is " $r^Tx = 0$ " since $Wz = z$ for every vector z in this plane. Every vector z not in this plane has a nonzero component Pz parallel to c pointing out of this plane; W reverses that component: $z = Pz + Qz$ so $Wz = WPz + WQz = -Pz + Qz$. For example take $r^T = \begin{bmatrix} -1 & 1 & 0 \end{bmatrix}$ and $c = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, so $W = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. This W is a permutation matrix that swaps the first two elements of a (row or column) vector. In general, every permutation of two elements is a reflector, leaving unchanged the plane of vectors with those elements equal, and reversing vectors with only those elements nonzero and of opposite sign. In most introductory texts, reflectors are all *orthogonal* reflectors with $c = (r^T)^T$ in an Euclidean space; but our reflectors work in a more general Affine space and can reverse any direction c not in the mirror-plane " $r^Tx = 0$ ". *Draw a picture.*

Why are Elementary Reflectors and Projectors called "Elementary" ?

Let $C := (c_1, c_2, \dots, c_k)$ be a "row" of k linearly independent vectors c_j drawn from an n-dimensional vector space; we assume $0 < k < n$. This C serves as a basis for a k-dimensional *subspace* S of the n-dimensional space. Many a basis $(C, \overline{C}) = (c_1, c_2, \dots, c_k, c_{k+1}, \dots, c_n)$ for the n-space can be built up from C by augmenting it successively by vectors c_k, c_{k+1}, …, c_n each chosen to be linearly independent of all its predecessors though otherwise arbitrary. Let R^T be the "column" of the first k linear functionals ("rows") $r^T{}_1$, $r^T{}_2$, …, $r^T{}_k$ in $(C, \overline{C})^{-1}$, whence it follows that $R^TC = I_k$, the k-by-k identity matrix, and $R^T\overline{C} = O_{k,n-k}$. (Can you see why?) Finally let $P := CR^T$. It is a projector because $O \neq P^2 = P \neq I$. (Why is $P \neq I$?) Complementary to P is the projector $Q := I - P$; and $W := I - 2P = 2Q - I$ is a reflector because $W^2 = I \neq W$. (What is the inverse of a reflector W ?) P and W leave unchanged the (n–k)-dimensional subspace $\overline{S}$ spanned by $\overline{C}$, which turns out to be also the plane of all vectors s that satisfy $R^Ts = o$. (Can you see why?)

In short, the n-space has been decomposed into a sum of two complementary subspaces S and $\overline{S}$ (neither of which determines the other because they need not be orthogonal), and projector P collapses the space into S along lines parallel to $\overline{S}$. Complementary projector Q collapses the space into $\overline{S}$ along lines parallel to S . Reflector W reflects the space in the mirror $\overline{S}$ along lines parallel to S . Although these operators preserve subspaces more general than the lines or hyperplanes preserved by elementary operators, these non-elementary operators can be decomposed into elementary operators, as we shall see now.

Let $P_j := c_jr^T{}_j$ for $j = 1, 2, \dots, k$. Evidently $P_j = P_j{}^2$ is an elementary projector; moreover $P_iP_j = O$ if $i \neq j$; can you see why? These elementary projectors are said to *annihilate* each other. Thus $P = P_1 + P_2 + \dots + P_k$ is the sum of elementary and mutually annihilating projectors. Similarly W is a product of k elementary reflectors $W_j := I - 2P_j$ in any order! (Confirm that $W_iW_j = W_jW_i$; they *commute*.) Finally we have a nontrivial theorem:

> Every projector can be decomposed into a sum of elementary and mutually annihilating projectors.
> Every reflector can be decomposed into a product of elementary and commuting reflectors.
> Every reflector or projector decomposes the space into complementary subspaces S and $\overline{S}$ as above.

This could be proved now, but its proof will become simpler after we have studied eigenvalues and eigenvectors.

Elementary Dilatators, Expansions and Compressions:

For any elementary projector $P := cr^T/r^Tc$ as defined above, and for any nonzero scalar μ, let $M(\mu) := I + (\mu-1)P = Q + \mu P$. The elementary reflector $W = M(-1)$ is a special case. In general, the *dilatator* $M(\mu)$ expands (if $|\mu| > 1$) or compresses (if $|\mu| < 1$) the space parallel to c while keeping the plane " $r^Tc = 0$ " unchanged. If $\mu < 0$ the dilatator also reflects the space in that unchanged plane. Other writers use "dilator", "dilation", "dilatation", "expander", "expansion", "compressor" or "compression" in place of "dilatator", or restrict them to Euclidean space with $c = (r^T)^T$ orthogonal (perpendicular) to the plane " $r^Tz = 0$ " left unchanged. Our c need not be orthogonal. *Draw pictures to illustrate diverse dilatators.*

Exercise: Confirm that $M(\mu)\,M(\lambda) = M(\mu\lambda)$ and therefore $M(\mu)^{-1} = M(1/\mu)$.

Example: For $r^T = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ and $c = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $M(\mu) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is the elementary operation that multiplies the second row or column of a 3-by-3 matrix by μ. Every elementary row operation that multiplies a row by a nonzero constant is a dilatator analogous to this one.

Elementary Shears:

Choose now any nonzero vector s and functional t^T constrained by $t^Ts = 0$. In other words choose first any plane " $t^Tx = 0$ " through o and then any nonzero vector s in that plane. Now an elementary shear $S^\mu := I + \mu st^T$ slides an arbitrary vector z to $y := S^\mu z$ by adding to z a multiple of s proportional to the distance between the two parallel planes " $t^Tx = 0$ " and " $t^Tx = t^Tz$ ", one through o and the other through z. *Draw pictures*.

A good way to visualize the effect of an elementary shear is to imagine a deck of playing cards stacked straight so that the stack's sides look like rectangles. Shearing the deck slides the cards, keeping their respective edges parallel, in such a way that the stack's sides remain parallelograms. The bottom card stays fixed in the plane " $t^Tx = 0$ ".

Exercise: Verify $S^\mu S^ß = S^{\mu+ß}$, and $(S^\mu)^{-1} = S^{-\mu}$.

Example: For $t^T = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ and $s = \begin{bmatrix} 6 \\ 0 \\ 7 \end{bmatrix}$, $S^\mu = I + \mu st^T = \begin{bmatrix} 1 & 6\mu & 0 \\ 0 & 1 & 0 \\ 0 & 7\mu & 1 \end{bmatrix}$ is the elementary row operation that adds 6μ times the second row to the first and 7μ times the second row to the third of a 3-by-3 matrix. Every elementary row operation that adds multiples of one row to others is analogously premultiplication by a shear. The inverse operation, subtracting, is a shear too. Analogous such elementary operations upon columns amount to postmultiplications by shears. Some writers restrict an elementary shear to have only one nonzero element in s and one in t^T, but this restriction gains nothing. What is common to all elementary operators? Each of them, of every kind, is determined by an operator of rank 1.

Conclusion:

Every nonsingular (invertible) square matrix is a product of Elementary Reflectors, Dilatators, and Shears. This is so because a product of these elementary row operations suffices to reduce the matrix to its Reduced Row Echelon Form, which must be the identity matrix I if the original matrix is invertible. Hence, the original matrix is a product of these operations' inverses, all of them elementary too. Similarly, every invertible linear operator mapping a finite-dimensional space to itself is a product of Elementary Reflectors, Dilatators, and Shears. This is so because it is true for the square nonsingular matrix that represents the linear operator in any basis for the space. *Work through the details. The following may help.*

REVIEW: Change of Basis as a Nonsingular Matrix:

Suppose $\mathbf{B}$ is one basis for a vector space, and $\mathbf{E}$ another. Since every basis vector $\mathbf{e}$ in $\mathbf{E}$ is expressible in terms of a column vector $c = \mathbf{B}^{-1}\mathbf{e}$, collecting those columns produces a matrix $C = \mathbf{B}^{-1}\mathbf{E}$ that appears in changes of coordinates as follows:

Let $\mathbf{x} = \mathbf{B}x$ be any vector in the space, and x its column vector of coordinates in the basis $\mathbf{B}$. Since $\mathbf{E}$ is a basis too, $\mathbf{x} = \mathbf{E}\mathbb{x}$ for some other column vector $\mathbb{x}$. The relation between x and $\mathbb{x}$ is this: $\mathbf{B}x = \mathbf{E}\mathbb{x}$, so $x = \mathbf{B}^{-1}\mathbf{E}\mathbb{x} = C\mathbb{x}$. And C^{-1} exists because $\mathbb{x} = \mathbf{E}^{-1}\mathbf{x} = \mathbf{E}^{-1}\mathbf{B}x = C^{-1}x$ is determinable from every x .

Conversely, any invertible matrix C is the matrix that changes coordinates to $x = \mathbf{B}^{-1}\mathbf{x}$ with one basis $\mathbf{B}$ from $\mathbb{x} = \mathbf{E}^{-1}\mathbf{x}$ with another basis $\mathbf{E} = \mathbf{B}C$; $x = C\mathbb{x}$.

Exercise: If $\mathbf{u}^T = u^T\mathbf{B}^{-1}$ in one basis $\mathbf{B}$, but $\mathbf{u}^T = \mathbb{u}^T\mathbf{E}^{-1}$ in another basis $\mathbf{E} = \mathbf{B}C$, how do we compute u^T given $\mathbb{u}^T$ and C ? How do we compute $\mathbf{u}^T\mathbf{x} = u^Tx$ given $\mathbb{u}^T$ and $\mathbb{x}$ and C ?

Now let $\mathbf{L}$ be any linear operator that maps the vector space to itself. This $\mathbf{L}$ is representable by a matrix L determined from a basis $\mathbf{B}$ of the vector space thus: if $\mathbf{x} = \mathbf{B}x$ maps to $\mathbf{y} = \mathbf{L}\mathbf{x}$ representable by column vector $y = \mathbf{B}^{-1}\mathbf{y}$, then $y = \mathbf{B}^{-1}\mathbf{L}\mathbf{x} = \mathbf{B}^{-1}\mathbf{L}\mathbf{B}x = Lx$ for the matrix $L := \mathbf{B}^{-1}\mathbf{L}\mathbf{B}$. A change of basis from $\mathbf{B}$ to $\mathbf{E} = \mathbf{B}C$ changes the matrix representing $\mathbf{L}$ from L to $\mathbb{L} := \mathbf{E}^{-1}\mathbf{L}\mathbf{E} = C^{-1}LC$ and the coordinates of $\mathbf{y}$ become $\mathbb{y} = \mathbb{L}\mathbb{x}$; can you confirm this?

Exercise: If linear operator $\mathbf{Q}$ maps a space with basis $\mathbf{B}$ linearly to another space with basis $\mathbf{H}$ of perhaps different dimension, what matrix Q represents $\mathbf{Q}$ in these bases?

Thus, a nonsingular matrix can represent a few things: a change of basis, an invertible map from a space to itself, an invertible map from one space to another of the same dimension, … . Out of context, a matrix does not say what it represents. The mathematically interesting question about it is this:

> Given a matrix that represents a linear map between two spaces, and given a characterization of the spaces but not the relevant basis or bases, what geometrical properties of the linear map can be inferred without knowing the basis or bases?