Holger Broman, August 19, 1999

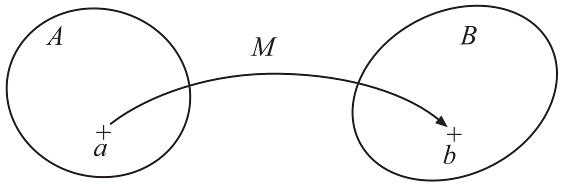
Norm-preserving linear mappings from $\mathbb{R}^N$ to $\mathbb{R}^N$

This section will give you an example of how to analyze a statement such as the one given in the heading. Let us reduce the complexity of the task by taking one concept at a time and finally synthesize to make the whole obvious.

A mapping M is denoted, for instance

$$M: \quad A \ni a \overset{M}{\curvearrowright} b \in B,$$

and illustrated by the Venn-diagram



The defining feature of a mapping is the uniqueness described by

$$\begin{aligned} & b = M(a) \text{ and } c = M(a) \\ \Rightarrow & \\ & c = b \end{aligned}$$

Note 1 A is a set containing elements such as a, and similarly for B and b. The set A is called the domain of M, and B the range of M (Swedish: definitionsmängd och värdeförråd). In the case of mappings from $\mathbb{R}^N$ to $\mathbb{R}^N$, $A = \mathbb{R}^N$ and $B \subseteq \mathbb{R}^N$.

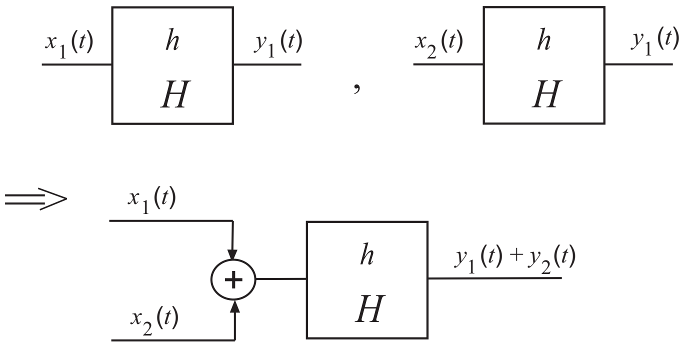
Note 2 The mapping is invertible iff

$$\begin{aligned} & b = M(a_1) \text{ and } b = M(a_2) \\ \Rightarrow & \\ & a_1 = a_2 \;\forall\; a_1,\, a_2 \end{aligned}$$

We then denote the inverse mapping by $a = M^{-1}(b)$.

———————— o O o ————————

Engineers call linear mappings by a different name; systems that obey the superposition principle. The illustration below shows one example.



One alternative engineering phrasing is to call it a linear system. In mathematical terminology:

> A mapping M is linear iff
>
> - $M(a_1 + a_2) = M(a_1) + M(a_2)$
>
> and
>
> - $M(c \cdot a) = cM(a), \ c \in R^1(C^1)$

Note 3 It is required that the addition $a_1 + a_2$ is defined and that the multiplication of a by a real (complex) number is defined. This is of course the case for mappings from $\mathbb{R}^N$ to $\mathbb{R}^N$.

——————— o O o ———————

Now to the problem of defining $\mathbb{R}^N$. As you all know, $\mathbb{R}^1$ is defined by

$$\mathbb{R}^1 = \left\{ \text{ all real numbers } \right\},$$

and

$$\mathbb{R}^2 = \left\{ \text{ all pairs of real numbers } \right\},$$

We thus define

$$\boxed{\mathbb{R}^N = \left\{ \text{ all } N\text{-tuples of real numbers } \right\}.}$$

Note 4 We often denote elements of $\mathbb{R}^N$ by bold characters, such as

$$\mathbf{x} = (x_1, x_2, \ldots, x_N), \quad x_n \in \mathbb{R}^1 \ \forall \ n$$

Note 5 There is an isomorphism between $\mathbb{R}^N$ and N-dimensional vectors with real scalars. In everyday speech we make extensive use of this isomorphism and do not differentiate between $\mathbb{R}^N$ and the set of N-dimensional vectors with real scalars.

Note 6 A polynomial is a scalar product. Define the notations

$$\mathbf{a}^T = (a_0 \; a_1 \; \cdots \; a_{N-1}),$$

and

$$\mathbf{x}^T = (x^0 \; x^1 \; \cdots \; x^{N-1}).$$

Then

$$p(x) \triangleq \mathbf{a}^T \mathbf{x} = \sum_{n=0}^{N-1} a_n \, x^n.$$

Thus, there is an isomorphism between $\mathbb{R}^N$ and polynomials, as defined by the coefficient vector $\mathbf{a}$.

————— o O o —————

What do we mean by a norm? An example is the area of a surface, another the volume of a body in three dimensions, a third the weight of a body. We generalize from these examples to require:

> A norm of the elements $\mathbf{x} \in \mathbb{C}^N$ is a mapping M to $\mathbb{R}^+$.
>
> $$M : \quad \mathbb{C}^N \ni \mathbf{x} \overset{M}{\curvearrowright} y \in \mathbb{R}^+$$
>
> It is required that
>
> - $M(\mathbf{x}) \geq 0$
> - $M(\mathbf{x}) = 0 \;\Rightarrow\; \mathbf{x} = 0$
> - $M(\mathbf{x}_1 + \mathbf{x}_2) \leq M(\mathbf{x}_1) + M(\mathbf{x}_2)$ – the triangle inequality
> - $M(c\mathbf{x}) = |c| M(\mathbf{x}), c \in \mathbb{C}^1$

————— o O o —————

We now start to synthesize. Let us start by a mapping M from $\mathbb{R}^N$ to $\mathbb{R}^N$:

$$M : \quad \mathbb{R}^N \ni \mathbf{x} \overset{M}{\curvearrowright} \mathbf{y} \in \mathbb{R}^N$$

What about a linear mapping? We require – among other things –

$$M(\mathbf{x}_1 + \mathbf{x}_2) = M(\mathbf{x}_1) + M(\mathbf{x}_2).$$

One obvious candidate is to let M be represented by a left multiplication by an N by N matrix A with real-valued entries (and $\mathbf{x}_1$, $\mathbf{x}_2$ by N-dimensional vectors):

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2.$$

We note with delight that

$$A\,c\,\mathbf{x} = c\,A\,\mathbf{x}, \quad c \in \mathbb{R}^1\,(\mathbb{C}^1),$$

so that this multiplication by a matrix indeed is a linear mapping. It is possible to prove that it is the only form of the mapping looked for, but we abstain from giving the proof. We summarize the result so far:

> A linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^N$ is equivalent to the multiplication of an N-dimensional real-valued vector by an N by N real-valued matrix.

——— o O o ———

Now we are closing in on the goal. Let us define the norm to use as the Euclidean norm $\| \mathbf{x} \|$ of a vector $\mathbf{x}$:

$$\| \mathbf{x} \|^2 = \mathbf{x}^T\mathbf{x} = \sum_{n=1}^{N} x_n^2$$

Note 7 You may use other norms – mappings from $\mathbb{R}^N$ to $\mathbb{R}^1$ – but we stay with the Euclidean norm, the length. Now, norm-preserving means that if

$$\mathbf{y} = A\mathbf{x},$$

then

$$\| \mathbf{y} \| = \| \mathbf{x} \|.$$

Let us do the math involved:

$$\| \mathbf{y} \|^2 = \| A\mathbf{x} \|^2 = \mathbf{x}^T A^T A\mathbf{x},$$

so that

$$\mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T\mathbf{x}.$$

Re-write

$$0 = \mathbf{x}^T A^T A\mathbf{x} - \mathbf{x}^T\mathbf{x} = \mathbf{x}^T(A^T A - I)\mathbf{x}.$$

For this to hold true <u>for all $\mathbf{x}$</u>, it is required that

$$A^T A = I.$$

Note 8 The matrix A must be such that

$$A^{-1} = A^T.$$

———————— o O o ————————

The final lap is now to come. What are the eigenvalues of A?

$$A\mathbf{g} = \lambda\mathbf{g}.$$

The fact that $\mathbf{g}$ may have complex-valued elements complicates matters somewhat, as does the fact that λ may be complex-valued. To comply with the requirement that a norm is non-negative, real valued, we must take the Hermitian transpose, i.e. transpose and complex conjugate:

$$\begin{aligned} & A\mathbf{g} = \lambda\mathbf{g} \\ \Rightarrow & \\ & \mathbf{g}^H A^H = \mathbf{g}^H \lambda^* = \lambda^* \mathbf{g}^H. \end{aligned}$$

As

$$A^H = A^T,$$

remember real-valued entries, we find

$$\mathbf{g}^H A^T A\mathbf{g} = \lambda^* \mathbf{g}^H \lambda \mathbf{g} = |\lambda|^2 \mathbf{g}^H \mathbf{g},$$

and finally, as $A^T A = I$,

$$\mathbf{g}^H \mathbf{g} = |\lambda|^2 \mathbf{g}^H \mathbf{g},$$

so that all eigenvalues of A must be on the unit circle.

———————— o O o ————————

We have reached the end of our journey:

> A norm-preserving linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^N$ is a real-valued matrix A such that
>
> $$A^T A = I.$$
>
> All eigenvalues of A have unit magnitude.

Note 9 The converse statement is true: any N by N real-valued matrix A that obeys $A^T A = I$ represents a norm-preserving linear mapping.

Note 10 If you write

$$A = [\mathbf{a}_1\ \mathbf{a}_2 \cdots \mathbf{a}_N],$$

then

$$I = A^T A = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_N^T \end{bmatrix} [\mathbf{a}_1 \cdots \mathbf{a}_N]$$

shows that the column vectors of A are orthonormal. Such a matrix A is called an orthogonal matrix (and not an orthonormal matrix).

Note 11 In the complex-valued case, using Hermitian transposition (transpose and complex conjugate), the equation

$$U^H U = I$$

defines a unitary matrix.

Note 12 The mapping defined by an orthogonal (unitary) matrix also preserves scalar products: Let

$$\mathbf{x}_1 = U\mathbf{x}, \quad \mathbf{y}_1 = U\mathbf{y},$$

then

$$\mathbf{x}_1^H \mathbf{y}_1 = \mathbf{x}^H U^H U \mathbf{y} = \mathbf{x}^H \mathbf{y}.$$

Please run the m-file normpres in Matlab.

```
% normpres.m
% in the 2D-case, a norm-preserving matrix is either a rotation or a reflection.
clear

alpha=pi/3;
c=cos(alpha);
s=sin(alpha);
Rot=[c -s ; s c]
Ref=[c s ; s -c] % Does NOT reflect in the line of slope alpha
x=rand(2,1)-[0.5;0.5];
xrot=Rot*x;
xref=Ref*x;
radius=norm(x);
t=0:0.01:2*pi;
figure(1), clf, hold on, axis equal
plot([0 x(1)],[0 x(2)],'r',x(1),x(2),'rp')
plot([0 xrot(1)],[0 xrot(2)],'g',xrot(1),xrot(2),'gp')
plot([0 xref(1)],[0 xref(2)],'b',xref(1),xref(2),'bp')
plot(radius*cos(t),radius*sin(t),'k')
plot(radius*[-1 1],radius*[-tan(alpha/2) tan(alpha/2)],'b--') % line of reflection
plot([x(1) xref(1)],[x(2) xref(2)],'k-.')

% let us take the projection while we are at it
Proj=(Ref+eye(2))/2
xproj=Proj*x;
plot([0 xproj(1)],[0 xproj(2)],'y',xproj(1),xproj(2),'yp')
title('Original-red, rotation pi/3 -green, reflection in pi/6 -blue, projection-yel

% Please find the eigenvalues of Rot, Ref, and Proj.
% Figure out why the are what they are, it is no accident.
% You may find it instructive to vary alpha, then do not trust
% the title of the figure.
% Run this program several times, random data!
```