Mathematical Tools

Per Åke Malmqvist

Department of Theoretical Chemistry
Chemical Center
Lund University

More accurately: A collection of stuff, of which some or the other may come in handy in most applied math.

What's in a name? Spaces and stuff.

A set of objects and a set of operations on these objects is called a space and the objects are called elements of the space.

Some spaces are defined in terms of the way their objects are composed of simpler objects, and how the operations are carried out in terms of operations on the simpler objects. This can be called a concrete space.

- **Example:** The set of real numbers, together with its arithmetic rules, is the space $\mathbb{R}$.

But **most** spaces are defined just by requiring that certain operations can be done on the elements, and that the results conform to some set of rules, axioms.

Such spaces are called abstract spaces.

- **Example:** A linear space.

Spaces and stuff (2).

The Hilbert space is an abstract space. This means that if one speaks of 'the Hilbert space', one is concerned only with properties that follow from the seven axioms that define such a space.

If one is dealing with some particular concrete space, one can say that it is 'a Hilbert space'. This means that it has operations conforming with the axioms, but there may well be lots of other properties which follow from the concrete realization.

- **Example:** The set of square-summable sequences of real numbers, $\mathbb{R}^\infty$, can in most respects be treated as a set of infinite-dimensioned vectors. Together with one additional axiom, this is a Hilbert space (called $l^2(\mathbb{R})$).

Classification of spaces: Size.

The simplest size concept is the cardinality of the set of elements:

$$\begin{aligned} \mathrm{card}(\{0,1,2\}) = \mathrm{card}(\{A,B,E\}) &= 3 \\ \mathrm{card}(\mathsf{N}) = \mathrm{card}(\mathsf{Z}) &= \aleph_0 \\ \mathrm{card}(\mathsf{R}) = \mathrm{card}([0,2\pi]) &= \aleph_1 \\ \mathrm{card}(\emptyset) &= 0 \end{aligned}$$

The cardinality is the number of possible different values.

Another very important concept is that of dimensionality:

$$\begin{aligned} \dim(\{0,1,2\}) &= 0 \\ \dim(\mathsf{R}) = 1 \\ \dim(\mathsf{R}^3) = 3 \\ \dim(\mathsf{C}) = 2 \\ \dim(\mathsf{C}^n) = 2n \end{aligned}$$

Classification of spaces: Size.

Another very important concept is that of dimensionality :

$$\begin{aligned} \dim(\{0,1,2\}) &= 0 \\ \dim(\mathsf{R}) = 1 \\ \dim(\mathsf{R}^3) = 3 \\ \dim(\mathsf{C}) = 2 \\ \dim(\mathsf{C}^n) = 2n \end{aligned}$$

The dimensionality is the number of continuously varying real parameters that are needed to specify an element of the space. In general:

$$\dim(S) = n$$

if in the neighborhood of any element of S, there is an invertible and continuous mapping $S \leftrightarrow \mathsf{R}^n$

Some non-trivial concepts: neighborhood? continuity? but for now, an intuitive understanding is enough.

Another classification of spaces: Operations.

Is a distance between any two elements defined? Then this is a normed space which is also called a metric space. See defs. 2.1 and 2.2. Examples:

Distance in R^3	Distance on the globe
Hamming distance	Typographic distance

Are linear scaling and adding defined? Then this is a linear space which is also called a vector space. See defs. 2.3. Examples:

Translations in R^3

The unnormalized electronic wave functions of pyrimidine

Is the space both linear **and** normed ? Then the distance function is specialized. If the space is also infinite-dimensional, it is a Banach space. See defs. 2.4. Example:

The maximum norm used in function fitting.

Is the space both linear and has a scalar product ? If the space is also infinite-dimensional, it is a Hilbert space. See defs. 2.5. Example:

The unnormalized electronic wave functions of pyrimidine

Little fleas have smaller fleas. . . ad infinitum

Given any linear space, the set of linear mappings between elements is in itself a linear space. If the original vector space has a norm or a scalar product, such properties can also be defined in the space of linear maps. The original space is called a carrier space while the space of mappings is called operator space.Its elements are called linear operators.

The operator space has one additional property, which maybe the carrier space did not have. In it, there is a multiplication defined, by composition of operators. A vector space where there is a multiplication rule, where the product is new elements, is also called a linear algebra . The space of mappings is called operator space .

Thus it is clear that any vector space also implies the existence of 'higher' vector spaces, namely the operator space, its operator space (which is called superoperator space), and so on. All these higher vector spaces are linear algebras: There is a multiplication rule defined.

Infinite spaces, closure and separability

For infinite spaces, some important properties are closure and separability.

Suppose we can show that for some infinite sequence of elements $x_1, x_2, \ldots$

$$\lim_{n,m\to\infty} \|x_m - x_n\| \;=\; 0$$

Question: Does this mean we can conclude that there is an element x such that

$$\lim_{n\to\infty} x_n \;= x$$

Then this is a complete space.

The property of completeness is also called closure. It is one of the properties required of a Banach or a Hilbert space (although we did not mention it before).

The other property is separability. For a vector space, this means that there are basis sets that are infinite but countable , which can be used to express any element in the space.

Some commonly used function spaces.

For infinite spaces, we will ignore the property of separability. (There exist non-separable vector spaces, but all the following spaces are in fact separable).

Notation	**Meaning**
$C^0(X)$	$\{f \quad \| \quad f$ is continuous in X $\}$
$C^l(X)$	$\{f \quad \| \quad \hat{D}^l f$ is continuous in X $(1 \leq l \leq \infty)\}$
$C^\infty(X)$	See above. Also called the class of smooth functions.
$D^l(X)$	As $C^l(X)$, with compact support
$C^l_B(X)$	As $C^l(X)$, but the value and derivatives are bounded.
$l^p(X)$	Banach space with norm $(\sum_X \|f(x)\|^p dx)^{1/p}$.
$L^p(X)$	Banach space with norm $(\int_X \|f(x)\|^p dx)^{1/p}$.
$H^1(X)$	Hilbert space with norm $\sqrt{\int_X (\|f(x)\|^2 + \|\nabla f(x)\|^2) dx}$.

The last examples, with $p = 2$, are particularly common. These are the Lebesgue $L^2(X)$spaces, and Sobolev $H^1(X)$ spaces which are Hilbert spaces, also called called the spaces of square-integrable functions or sequences, and for Sobolev, also with square-integrable derivatives.

An example of a Hilbert space: $L^2(\mathbb{R})$.

$L^2(\mathbb{X})$ is a Hilbert space, because it has a scalar product. With $\mathbb{X} = \mathbb{R}$,

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)^* g(x) dx$$

A typical element is $f(x) = \mathrm{e}^{-x^2}$. Define two typical operators $\hat{A}$ and $\hat{B}$:

$$\hat{A} = (2x - \partial/\partial x) \quad i.e. \quad \left[\hat{A} f\right](x) = 2x f(x) - f'(x)$$
$$\hat{B} = (2x + \partial/\partial x) \quad i.e. \quad \left[\hat{B} f\right](x) = 2x f(x) + f'(x)$$

Then

$$\begin{aligned} \hat{A}\hat{B} &= 4x^2 - \partial^2/\partial x^2 - 2 \\ \hat{B}\hat{A} &= 4x^2 - \partial^2/\partial x^2 + 2 \end{aligned}$$

We see that the two operators do not commute: $\hat{A}\hat{B} \neq \hat{B}\hat{A}$.

Basis representation

Assume a vector space S, with dimension n. We note the following facts.

A. Any set of m vectors $\{x_i\}_{i=1}^m$ defines a subspace S' of S, denoted

$$S' = \mathrm{span}(\{x_1, \ldots, x_m\})$$

and defined as the set of all possible vectors of the form

$$x = \alpha_1 x_1 + \cdots + \alpha_m x_m, \quad \alpha_i \in \mathsf{R}$$

B. The vectors in this subspace can be represented by the mapping

$$(\alpha_1, \ldots \alpha_m) \mapsto x$$

Scaling: $(\alpha\alpha_1, \ldots \alpha\alpha_m) \mapsto \alpha x$
Adding: If $(\beta_1, \ldots \beta_m) \mapsto y$, then

$$(\alpha_1 + \beta_1, \ldots \alpha_m + \beta_m) \mapsto x + y$$

This mapping is generally "many-to-one".

Basis set and independence.

C. If

$$(\alpha_1, \ldots \alpha_m) \mapsto 0 \Rightarrow (\forall i : \alpha_i = 0)$$

then the vector set is linearly independent and then $\dim(S') = m$, the mapping is 1:1, and the set $\{x_i\}_{i=1}^m$ is a basis set of S'. Else, the set is linearly dependent and then $\dim(S') < m$.

D. If the vector set $\{x_i\}_{i=1}^m$ is linearly dependent, it is always possible to discard one or more of the vectors, just keeping $m' < m$ vectors relabelled as $\{x'_1, \ldots, x'_{m'}\}$, where $m' = \dim(S')$ and $\{x'_1, \ldots, x'_{m'}\}$ is a basis set for S'.

This is simplest to do if the vector space has a scalar product. A famous procedure is the Gram-Schmidt orthonormalization, which actually produces an orthonormal basis with the same span as the original vectors.

The Gram-Schmidt procedure

E. If there is a scalar product, the following procedure will create an orthonormal basis:

```
n:=0
for k=1 to m
   v := x_k - sum( b_i <b_i|x_k> ,i=1..n)
   if ( |v|>0 ) then
     n:=n+1
     b_{n} := v/|v|
   end if
end loop
```

The input data are the m vectors $\{\mathbf{x}_k\}_{k=1}^{m}$, and the output consists of the n vectors $\{\mathbf{b}_i\}_{i=1}^{n}$. The $\mathbf{b}$-vectors are orthonormal, and span the same space as the $\mathbf{x}$-vectors.

Sequences, arrays, matrices

An ordered set of numbers is called a sequence or an array. If scaling and elementwise addition are meaningful operations, they are regarded as vectors. Indeed, they are vectors, in the vector space called $\mathbb{R}^n$ (or $\mathbb{C}^n$), where n (or $2n$) is the dimension. If a basis set is defined in any particular other vector space, the representation of these vectors by the array of expansion coefficients shows that every space with a basis set representation is in some sense equivalent with any other space with the same dimension.

The most common way of writing vectors and linear operations in a finite-dimensional space is thus by means of 'column vectors' and matrices.

Matrix formalism and basis sets.

Matrix formalism in an n-dimensional vector space $\mathcal{S}$ works as follows:

A vector in $\mathcal{S}$ is written as a column vector (i.e. an $n \times 1$ matrix) with expansion coefficients, while a vector in the dual space $\mathcal{S}^*$ is written as a row vector (i.e. an $1 \times n$ matrix). The basis set is written as a row of symbols for the basis elements, where f is an element of the vector space being studied, $\{\chi_k\}_{k=1}^n$ is a basis, and $\{x_k\}_{k=1}^n$ are the expansion coefficients.

$$f = (\chi_1, \chi_2, \ldots, \chi_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Matrix formalism (2)

Assume a scalar product space and an orthonormal basis. In that case, the hermitian conjugate is meaningful, and is represented by the transposed complex conjugate of the matrix:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^\dagger \stackrel{\text{def}}{=} (x_1^*, x_2^*, \ldots, x_n^*)$$

The definition of equality for linear operations in a vector space is that they give identical results when applied to any vector: if for any given operator $\hat{A}$, there is a matrix $\mathbf{A}$ such that

$$f = (\chi_1, \chi_2, \ldots, \chi_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow \hat{A} f = (\chi_1, \chi_2, \ldots, \chi_n) \mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

is true for any vector f, then $\mathbf{A}$ is a matrix representation of $\hat{A}$.

Matrix formalism (3)

By chosing the matrix elements as

$$A_{kl} = \langle \chi_k | \hat{A} \chi_l \rangle$$

the previous equation is correct for the special case that all the expansion coefficients are zero except one, which is $= 1$. By linearity, it is then true for any vector. So every linear operator has a matrix representation in a given basis, and when the basis is orthonormal, it is given above.

Assume the operators $\hat{A}$ and $\hat{B}$ are represented by the matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. Some properties of matrix representations of operations are then:

$$\begin{aligned} \hat{1} &\quad \text{is represented by} \quad \mathbf{1} \\ \hat{A}\hat{B} &\quad \text{is represented by} \quad \mathbf{AB} \\ \hat{A}^{-1} &\quad \text{is represented by} \quad \mathbf{A}^{-1} \\ \hat{A}^{n} &\quad \text{is represented by} \quad \mathbf{A}^{n} \end{aligned}$$

Changing the basis – basis set transformation

Any two bases are related by an invertible linear transformation, in the form of multiplication with a nonsingular transformation matrix, say $\mathbf{T}$:

$$(|\chi_1\rangle, |\chi_2\rangle, \ldots, |\chi_n\rangle) = (|\eta_1\rangle, |\eta_2\rangle, \ldots, |\eta_n\rangle) \begin{pmatrix} T_{11} & T_{12} & \ldots & T_{1n} \\ T_{21} & T_{22} & \ldots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & \ldots & \ldots & T_{nn} \end{pmatrix}$$

Expand a function using two different basis sets:

$$|f\rangle = (|\chi_1\rangle, |\chi_2\rangle, \ldots, |\chi_n\rangle) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (|\eta_1\rangle, |\eta_2\rangle, \ldots, |\eta_n\rangle) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Changing the basis (2)

Substitute:

$$|f\rangle = (|\eta_1\rangle, |\eta_2\rangle, \ldots, |\eta_n\rangle)\ \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (|\eta_1\rangle, |\eta_2\rangle, \ldots, |\eta_n\rangle) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

The relation between the representation vectors is then simply

$$\mathbf{Tx} = \mathbf{y}, \qquad \mathbf{x} = \mathbf{T}^{-1}\mathbf{y}$$

and the matrix representations of an operator transforms as

$$\mathbf{A}^{(\chi)} = \mathbf{T}^{-1}\mathbf{A}^{(\eta)}\mathbf{T},$$

called a *similarity transformation*.

Representation independence

A number of properties of $n \times n$ matrices are invariant to similarity transformations. So in a basis set representation of an operator, the matrix depends on the basis used, some properties do not change with the basis – they are basis set independent. Some such properties are:

$$\mathrm{tr}(\mathbf{A}) = \mathrm{sp}(\mathbf{A}) \stackrel{\mathrm{def}}{=} \sum_i A_{ii}$$

$$\det(\mathbf{A})$$

$$\mathrm{rank}(\mathbf{A})$$

$$\rho_2(\mathbf{A}) \quad \text{(spectral radius)}$$

$$\mathrm{tr}(f(\mathbf{A})) \quad \text{(where } f \text{ is an analytic function)}$$

$$\det(f(\mathbf{A}))$$

Trace of a matrix

The trace of a square matrix is the sum of its diagonal entries. If the matrix is a product of two matrices, the trace does not depend on the order in which the product is computed (not even if the factors are rectangular!):

$$\mathrm{tr}(\mathbf{AB}) = \sum_{i=1}^{n}(\sum_{k=1}^{m} A_{ik}B_{ki}) = \sum_{k=1}^{m}(\sum_{i=1}^{n} B_{ki}A_{ik}) = \mathrm{tr}(\mathbf{BA})$$

The trace of a product of three or more matrices is unchanged if the order is cyclically permuted:

$$\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{A}(\mathbf{BC})) = \mathrm{tr}((\mathbf{BC})\mathbf{A}) = \mathrm{tr}(\mathbf{BCA})$$

The properties are extended from matrices to operators by using basis set representations.

Determinant of a matrix

The determinant of a square matrix is defined by the properties that: (a), $\det(\mathbf{1}) = 1$, (b), if any row is multiplied with a constant, the determinant is scaled with this constant, and (c), if any row is added to another row, the determinant is unchanged.

Some important properties are e.g. that $\det(\mathbf{A}^T) = \det(\mathbf{A})$ so any true statement about the determinant is true if rows are changed to columns and vice versa. Also:

- $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.
- Therefore, $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.
- Therefore, the matrix inverse exists if, and only if, its determinant is non-zero. Such a matrix is called non-singular. If the determinant is zero, it is singular.

Usually, the determinant function cannot be computed for operators in infinite spaces.

Representation independence (2)

If the inverse of a (linear) operator (e.g. a square matrix) exists, it is called non-singular; else it is singular. A product is singular if any of its factors is singular.

The spectrum of a matrix or an operator is the set of values

$$\{ z \,:\, (\mathbf{A} - z\mathbf{1}) \quad \text{is singular.} \}$$

Since $(\mathbf{T}^{-1}\mathbf{A}\mathbf{T} - z\mathbf{1}) = \mathbf{T}^{-1}(\mathbf{A} - z\mathbf{1})\mathbf{T}$ is singular when $(\mathbf{A} - z\mathbf{1})$ is singular, all functions of the spectrum are representation invariant.

Also the reverse is true: Any property that is representation independent is a function of the spectrum.

Any value z belonging to the spectrum is an eigenvalue, with (at least) one non-zero vector called an eigenvector $\mathbf{v}$: Since $(\mathbf{A} - z\mathbf{1})$ is singular, there is such a vector for which $(\mathbf{A} - z\mathbf{1})\mathbf{v} = \mathbf{0}$, i.e.,

$$\mathbf{A}\mathbf{v} = z\mathbf{v}$$

Eigenvalues of Symmetric real two by two matrices

The two-by-two matrices are so easy that the eigensystem can be obtained generally:

$$\det\begin{pmatrix} A_{11}-z & A_{12} \\ A_{12} & A_{22}-z \end{pmatrix} = z^2 - (A_{11}+A_{22})z + (A_{11}A_{22} - A_{12}^2) = 0$$

$$z = \frac{A_{11}+A_{22}}{2} \pm \sqrt{\left(\frac{A_{11}-A_{22}}{2}\right)^2 + A_{12}^2}$$

The eigenvalues are centered around $\frac{A_{11}+A_{22}}{2}$.

They are spaced out from this value by at least $\pm\frac{A_{11}-A_{22}}{2}$, and by at least $\pm A_{12}$.

Their sum is $= A_{11} + A_{22} = \mathrm{tr}(\mathbf{A})$, Their product is $= A_{11}A_{22} - A_{12}^2 = \det(\mathbf{A})$.

These last two properties are true for any (diagonalizable) matrix: the trace and product are the trace and the determinant of the matrix!