1. Sets, assignments, lists, and maps

The basic objects of Mathematics are sets and maps. Linear Algebra is perhaps the first course where this fact becomes evident and where it can be illustrated in a relative straightforward context. Since a complete understanding of the course material requires a thorough appreciation of the basic facts about maps, we begin with these and their simpler cousins, lists and assignments, after a brief review of standard language and notation concerning sets.

Sets

Sets of interest in these notes include

- the **natural numbers** : $\mathbb{N} := \{1, 2, \ldots\}$;
- the **integers** : $\mathbb{Z} := \{\ldots, -1, 0, 1, \ldots\} = (-\mathbb{N}) \cup \{0\} \cup \mathbb{N}$;
- the nonnegative integers : $\mathbb{Z}_+ := \{p \in \mathbb{Z} : p \geq 0\}$;
- the **rational numbers** : $\mathbb{Z} \div \mathbb{N} := \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}$;
- the **real numbers** and the nonnegative reals : $\mathbb{R}$, $\quad \mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$;
- the **complex numbers** : $\mathbb{C} := \mathbb{R} + \mathrm{i}\mathbb{R} = \{x + \mathrm{i}y : x, y \in \mathbb{R}\}$, $\quad \mathrm{i} := \sqrt{-1}$.

As these examples show, a set is often specified in the form $\{x : P(x)\}$ which is read 'the set of all x that have the property $P(x)$'. Note the use of the colon, ':', (rather than a vertical bar, '|') to separate, the initial, provisional, description of the typical element of the set, from the conditions imposed on it for membership in the set. In these notes, braces, '{', '}', are used solely in the description of sets.

Standard notation concerning sets includes:

- $\#S$ denotes the **cardinality** of the set S, i.e., the count of its elements.
- $x \in S$ and $S \ni x$ both mean that x is an element of S.
- $S \subset T$, $T \supset S$ both mean that S is a **subset** of T, i.e., all the elements of S are also elements of T; if we want to convey that S is a **proper subset** of T, meaning that $S \subset T$ but $S \neq T$, we write $S \subsetneq T$.
- $\{\}$ denotes the **empty set**, the set with no elements.
- $S \cap T := \{x : x \in S \text{ and } x \in T\}$ is the **intersection** of S and T.
- $S \cup T := \{x : x \in S \text{ or } x \in T\}$ is the **union** of S and T.
- $S \backslash T := \{x : x \in S \text{ but not } x \in T\}$ is the **difference** of S from T and is often read 'S take away T'. In these notes, this difference is *never* written $S - T$, as the latter is reserved for the set $\{s - t : s \in S, t \in T\}$ formable when both S and T are subsets of the same vector space.

1.1 What is the standard name for the elements of $\mathbb{R}\backslash(\mathbb{Z} \div \mathbb{N})$?

1.2 What is the standard name for the elements of $\mathrm{i}\mathbb{R}$?

1.3 Work out each of the following sets. (a) $(\{-1, 0, 1\} \cap \mathbb{N}) \cup \{-2\}$; (b) $(\{-1, 0, 1\} \cup \{-2\}) \cap \mathbb{N}$; (c) $\mathbb{Z}\backslash(2\mathbb{Z})$; (d) $\{z^2 : z \in \mathrm{i}\mathbb{R}\}$.

1.4 Determine $\#((\mathbb{R}_+\backslash\{x \in \mathbb{R} : x^2 > 16\}) \cap \mathbb{N})$.

Assignments

Definition: An **assignment** or, more precisely, an **assignment on** I or I**-assignment**

$$f = (f_i)_{i\in I} = (f_i : i \in I)$$

associates with each element i in its **domain** (or, **index set**)

$$\operatorname{dom} f := I$$

some **term** or **item** or **entry** or **value** f_i. In symbols:

$$f : \operatorname{dom} f : i \mapsto f_i.$$

The set

$$\operatorname{ran} f := \{f_i : i \in \operatorname{dom} f\}$$

of all items appearing in the assignment f is called the **range** of the assignment.

If also g is an assignment, then $f = g$ exactly when $f_i = g_i$ for all $i \in \operatorname{dom} f = \operatorname{dom} g$.

Very confusingly, many mathematicians call an assignment an *indexed set*, even though it is most certainly not a set. The term **family** is also used; however it, too, smacks too much of a set or collection.

We call the assignment f **1-1** if $f_i = f_j \Longrightarrow i = j$.

The simplest assignment is the **empty assignment**, (), i.e., the unique assignment whose domain is the empty set. Note that the empty assignment is 1-1 (why??).

An assignment with domain the set

$$\underline{n} := \{1, 2, \ldots, n\}$$

of the first n natural numbers is called a **list**, or, more explicitly, an n**-list**.

To specify an n-list f, it is sufficient to list its terms or values:

$$f = (f_1, f_2, \ldots, f_n).$$

For example, the **cartesian product**

$$\times_{i=1}^n X_i := X_1 \times X_2 \times \cdots \times X_n := \{(x_1, x_2, \ldots, x_n) : x_i \in X_i, i = 1{:}n\}$$

of the set sequence $X_1, \ldots, X_n$ is, by definition, the collection of all n-lists with the ith item or **coordinate** taken from X_i, all i.

In these notes, we deal with n*-vectors*, i.e., n-lists of *numbers*, such as the 3-lists $(1, 3.14, -14)$ or $(3, 3, 3)$. (Note that the *list* $(3, 3, 3)$ is quite different from the *set* $\{3, 3, 3\}$. The list $(3, 3, 3)$ has three terms, while the set $\{3, 3, 3\}$ has exactly one element.)

Definition: An n**-vector** is a list of n scalars (numbers). The collection of all **real** (**complex**) n-vectors is denoted by $\mathbb{R}^n$ ($\mathbb{C}^n$).

In MATLAB, there are (at least) two ways to specify an n-vector, namely as a one-row matrix (colloquially known as a **row vector**), or as a one-column matrix (colloquially known as a **column vector**). For example, we can record the 3-vector $x = (1.3, 3.14, -15)$ as the one-row matrix

```
x_as_row = [1.3,3.14,-15];
```

or as the one-column matrix

```
x_as_col = [1.3;3.14;-15];
```

One can also write a one-column matrix as a column, without the need for the semicolons, e.g.,

```
x_as_col = [1.3
     3.14
   -15];
```

□

Back to general assignments. If $\operatorname{dom} f$ is finite, say $\#\operatorname{dom} f = n$, then we could always describe f by listing the n pairs (i, f_i), $i \in \operatorname{dom} f$, in some fashion. However, that may not always be the most helpful thing to do. Here is a famous example.

During the Cholera outbreak in 1854 in London, Dr. John Snow recorded the deaths by address, thus setting up an assignment whose domain consisted of all the houses in London. But he did not simply make a list of all the addresses and then record the deaths in that list. Rather, he took a map of London and marked the number of deaths at each address right on the map (not bothering to record the value 0 of no death). He found that the deaths clustered around one particular public water pump, jumped to a conclusion (remember that this was well before Pasteur's discoveries), had the handle of that pump removed and had the satisfaction of seeing the epidemic fade.

Thus, one way to think of an assignment is to visualize its domain in some convenient fashion, and, 'at' each element of the domain, its assigned item or value.

This is routinely done for matrices, another basic object in these notes.

1.5 In some courses, students are assigned to specific seats in the class room. (a) If you were the instructor in such a class, how would you record this seating assignment? (b) What are the range and domain of this assignment?

1.6 A **relation** between the sets X and Y is any subset of $X \times Y$. Each such relation relates or associates with some elements of X one or more elements of Y. For each of the following relations, determine whether or not it provides an assignment on the set $X := \underline{3} =: Y$. (i) $R = X \times Y$; (ii) $R = \{(x, x) : x \in X\}$; (iii) $R = \{(1, 2), (2, 2)\}$; (iv) $R = \{(1, 2), (2, 1)\}$; (v) $R = \{(1, 2), (3, 1), (2, 1)\}$; (vi) $R = \{(1, 2), (2, 2), (3, 1), (2, 1)\}$.

Matrices

Definition: A **matrix**, or, more precisely, an $m \times n$**-matrix**, is any assignment with domain the cartesian product

$$\underline{m} \times \underline{n} \;=\; \{(i, j) : i \in \underline{m}, j \in \underline{n}\}$$

of $\underline{m}$ with $\underline{n}$, for some nonnegative m and n.

The collection of all **real**, resp. **complex** $m \times n$-matrices is denoted by $\mathbb{R}^{m\times n}$, resp. $\mathbb{C}^{m\times n}$.

In other words, a matrix has a rectangular domain. Correspondingly, it is customary to display such an $m \times n$-matrix A as a rectangle of items:

$$A \;=\; \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}$$

rather than as a list of pairs. This means that we must think of its domain rotated clockwise 90° when compared to the ordinary $(x, y)-$plane, i.e., the domain of many other bivariate assignments (or maps).

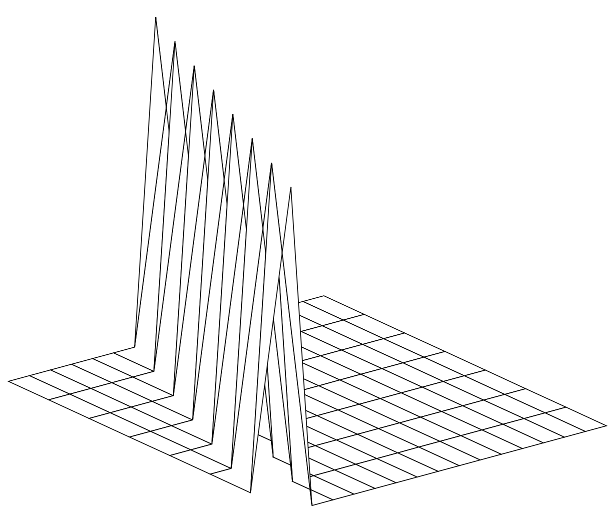
This way of displaying a matrix has led to the following language.

Let A be an $m \times n$-matrix. The item
$$A_{i,j} =: A(i,j)$$
corresponding to the index (i, j) is also called the **(i, j)-entry** of A. The list $A(i, :) := (A(i, j) : j \in \underline{n})$ is called the **ith row of** A, the list $A(:, j) := (A(i, j) : i \in \underline{m})$ is called the **jth column of** A, and the list $(A(i, i) : 1 \le i \le \min\{m, n\})$ is called the **(main) diagonal of** A. By definition, $A^{\rm t}$ denotes the **transpose** of the matrix A, i.e., the $n \times m$-matrix whose (i, j)-entry is $A(j, i)$, all i, j. Because of its importance in the later parts of these notes, we usually use the **conjugate transpose** $A^{\rm c} := \overline{A}^{\rm t}$ whose (i, j)-entry is the scalar $\overline{A(j, i)}$, with $\overline{\alpha}$ the complex conjugate of the scalar α.

When $m = n$, A is called a **square matrix** of **order** n.

The notation $A(i, :)$ for the ith row and $A(:, j)$ for the jth column of the matrix A is taken from `MATLAB`, where, however, `A(i,:)` is a one-row matrix and `A(:,j)` is a one-column matrix (rather than just a vector). The (main) diagonal of a matrix `A` is obtained in `MATLAB` by the command `diag(A)`, which returns, in a one-column matrix, the list of the diagonal elements. The conjugate transpose of a matrix `A` is obtained by `A'`. This is the same as the transpose if `A` is real. To get the mere *transpose* $\mathtt{A}^{\rm t}$ in the contrary case, you must use the notation `A.'` which is strange since there is nothing *pointwise* about this operation.

The above-mentioned need to look at displays of matrices sideways is further compounded when we use `MATLAB` to plot a matrix. Here, for example, is the 'picture' of the 8×16-matrix $A :=$ `eye(8,16)` as generated by the command `mesh(eye(8,16))`. This matrix has all its diagonal entries equal to 1 and all other entries equal to 0. But note that a careless interpretation of this figure would lead one to see a matrix with 16 rows and only 8 columns, due to the fact that `MATLAB`'s `mesh(A)` command interprets `A(i,j)` as the value of a bivariate function at the point `(j,i)`.



The rectangular identity matrix `eye(8,16)` as plotted in `MATLAB`

□

While lists can be concatenated in just one way, by letting one follow the other, matrices can be 'concatenated' by laying them next to each other and/or one underneath the other. The only requirement is that the result be again a matrix. If, for example,

$$A := [1 \quad 2], \qquad B := \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \qquad C := \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix},$$

then there are four different ways to 'concatenate' these three matrices, namely

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \begin{bmatrix} 4 & 5 & 3 \\ 7 & 8 & 6 \\ 1 & 2 & 9 \end{bmatrix}, \begin{bmatrix} 3 & 1 & 2 \\ 6 & 4 & 5 \\ 9 & 7 & 8 \end{bmatrix}, \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 1 & 2 \end{bmatrix}.$$

In `MATLAB`, one would write the three matrices

```
A = [1 2];   B = [3;6;9];   C = [4 5; 7 8];
```

and would describe the four possible 'concatenations'

```
[[A;C],B];   [[C;A],B];   [B,[A;C]];   [B,[C;A]];
```

We saw earlier that even vectors are described in `MATLAB` by matrices since `MATLAB` only knows matrices.

□

1.7 For the matrix A given by `[[0 0 0 0];eye(2,4)]`, determine the following items: (a) the main diagonal; (b) the second column; (c) the third row; (d) $A(3,2)$; (e) $A^{\rm t}$; (f) $A^{\rm c}$.

Lists of lists

Matrices are often used to record or represent a list $f = (f_1, f_2, \ldots, f_n)$ in which all the items f_j are themselves lists. This can always be done if all the items f_j in that list have the same length, i.e., for some m and all j, $\#f_j = m$. Further, it can be done in two ways, by columns or by rows.

Offhand, it seems more natural to think of a matrix as a list of its rows, particularly since we are used to writing things from left to right. Nevertheless, in these notes, it will always be done by columns, i.e., the sequence $(f_1, f_2, \ldots, f_n)$ of m-vectors will be associated with the $m \times n$-matrix A whose jth column is f_j, all j. We write this fact in this way:

$$A = [f_1, f_2, \ldots, f_n]; \qquad \text{i.e., } A(:,j) = f_j,\ j = 1{:}n.$$

This makes it acceptable to denote by

$$\#A$$

the number of columns of the matrix A. If I need to refer to the number of rows of A, I will simply count the number of columns of its transpose, $A^{\rm t}$, or its conjugate transpose, $A^{\rm c}$, i.e., write

$$\#A^{\rm t} \quad \text{or} \quad \#A^{\rm c},$$

rather than introduce yet another notation.

Here is a picturesque example of a list of lists, concerning the plotting of a **polyhedron**, specifically the regular **octahedron**. Its vertex set consists of the three unit vectors and their negatives, i.e.:

```
vs = [1 0 0; -1 0 0; 0 1 0; 0 -1 0; 0 0 1; 0 0 -1]';
```

Each face is a triangle, and we specify it here by giving the index, in the vertex array vs, of each of its three vertices:

```
ff = [2 4 5; 2 5 3; 4 1 5; 2 6 4]';
bf = [2 3 6; 6 1 4; 6 3 1; 5 1 3]';
```

The faces have been organized into front faces and back faces, in anticipation of the plotting about to be done, in which we want to plot the front faces strongly, but only lightly indicate the back faces. Be sure to look for specific faces in the figure below, in which the six vertices are numbered as in vs. E.g., the first front face, specified by the first column of ff, involves the vertices numbered 2, 4, 5; it is the face we are viewing head-on.

First, we set the frame:

```
axis([-1 1 -1 1 -1 1])
hold on, axis off
```

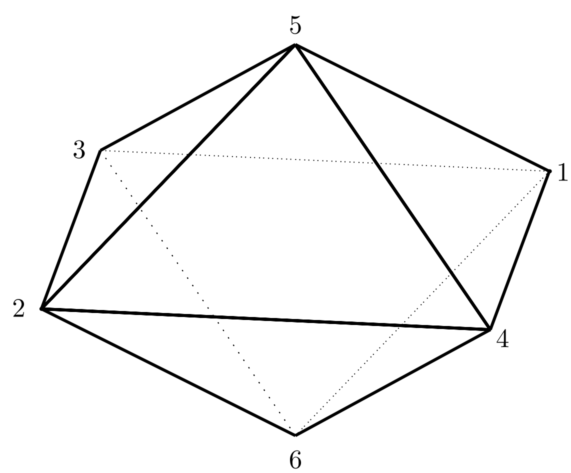
Then we plot the back-faces first (using r=[1 2 3 1] to make sure that we plot closed triangles):

```
r = [1 2 3 1];
for j=1:4
   plot3(vs(1,bf(r,j)),vs(2,bf(r,j)),vs(3,bf(r,j)),':')
end
```

Then, finally, we plot the front faces and finish the picture:

```
for j=1:4
   plot3(vs(1,ff(r,j)),vs(2,ff(r,j)),vs(3,ff(r,j)),'linew',1.5);
end
hold off
```

Here is the resulting figure (obtained by the command print -deps2 figoctah.eps which generates a postscript file). I have labeled all the vertices by their index in the vertex list vs.



The regular octahedron.

1.8 The regular octahedron is one of five regular solids. Write a MATLAB function drawrs(n) that will, for input $n \in (1, 2, 3, 4, 5)$, draw the regular (tetrahedron, cube, octahedron, dodecahedron, icosahedron).

□

Maps

Definition: A **map**

$$f : X \to Y : x \mapsto f(x)$$

associates with each element x of its **domain** $\operatorname{dom} f := X$ a unique element $y = f(x)$, called the **value of f at x**, from its **target** $\operatorname{tar} f := Y$. If g is also a map, then $f = g$ means that $\operatorname{dom} f = \operatorname{dom} g$, $\operatorname{tar} f = \operatorname{tar} g$, and $f(x) = g(x)$ for all $x \in \operatorname{dom} f$.

The collection

$$\operatorname{ran} f := \{f(x) : x \in X\}$$

of all values taken by f is called the **range of f**. More generally, for any subset Z of X,

$$fZ := f(Z) := \{f(z) : z \in Z\}$$

is called the **image of Z under f**. In these terms,

$$\operatorname{ran} f = f(\operatorname{dom} f).$$

Also, for any $U \subset Y$, the set

$$f^{-1}U := \{x \in X : f(x) \in U\}$$

is called the **pre-image of U under f**. The collection of all maps from X to Y is denoted by

$$Y^X \quad \text{or} \quad (X \to Y).$$

Names other than **map** are in use, such as **mapping**, **operator**, **morphism**, **transformation** etc., all longer than 'map'. A scalar-valued map is often called a **function**. Somewhat confusingly, many mathematicians use the term 'range' for what we have called here 'target'; the same mathematicians use the term **image** for what we have called here 'range'.

Every map $f : X \to Y$ gives rise to an assignment on X, namely the assignment $(f(x) : x \in X)$. On the other hand, an assignment f on X gives rise to *many* maps, one for each Y that contains $\operatorname{ran} f$, by the prescription $X \to Y : x \mapsto f_x$. We call this the **map into Y given by the assignment f**.

If X is empty, then Y^X consists of exactly one element, namely the map given by the empty assignment, and this even holds if Y is empty.

However, if Y is empty and X is not, then there can be no map from X to Y, since any such map would have to associate with each $x \in X$ some $y \in Y$, yet there are no $y \in Y$ to associate with.

"Wait a minute!", you now say, "How did we manage when X was empty?" Well, if X is empty, then there is no $x \in X$, hence the question of what element of Y to associate with never comes up. Isn't Mathematics slick?

1.9 Which of the following lists of pairs describes a map from {o,u,i,a} to {t,h,s}? A: ((u,s), (i,s), (a,t), (o,h), (i,s)); B: ((i,t), (a,s), (o,h), (i,s), (u,s)); C: ((a,s), (i,t), (u,h), (a,s), (i,t)).

1.10 For each of the following `MATLAB` maps, determine their range, as maps on real 2-by-3 matrices: (a) `A` $\mapsto$ `max(A)`; (b) `A` $\mapsto$ `A(:,2)`; (c) `A` $\mapsto$ `diag(A)`; (d) `A` $\mapsto$ `size(A)`; (e) `A` $\mapsto$ `length(A)`; (f) `A` $\mapsto$ `cos(A)`; (g) `A` $\mapsto$ `ones(A)`; (h) `A` $\mapsto$ `sum(A)`.

1.11 The **characteristic function** χ_S of the subset S of the set T is, by definition, the function on T that is 1 on S and 0 otherwise:

$$\chi_S : T \to \{0,1\} : t \mapsto \begin{cases} 1, & \text{if } t \in S; \\ 0, & \text{otherwise.} \end{cases}$$

Let R and S be subsets of T. Prove that (a) $\chi_{R\cup S} = \max(\chi_R, \chi_S)$; (b) $\chi_{R\cap S} = \min(\chi_R, \chi_S) = \chi_R \chi_S$; (c) $\chi_{R\backslash S} = \chi_R(1-\chi_S)$. (d) $R \subset S$ iff $\chi_R \le \chi_S$.

1.12 Let $f : T \to U$, and consider the map from subsets of U to subsets of T given by the rule

$$R \mapsto f^{-1}R := \{t \in T : f(t) \in R\}.$$

Prove that this map commutes with the set operations of union, intersection and 'take away', i.e., for any subsets R and S of U, (a) $f^{-1}(R \cup S) = (f^{-1}R) \cup (f^{-1}S)$; (b) $f^{-1}(R \cap S) = (f^{-1}R) \cap (f^{-1}S)$; (c) $f^{-1}(R \backslash S) = (f^{-1}R) \backslash (f^{-1}S)$.

1-1 and onto

In effect, a map is an assignment together with a target, with the target necessarily containing the range of the assignment. A major reason for introducing the concept of *map* (as distinct from the notion of *assignment*) is in order to raise the following basic question:

Given the map $f : X \to Y$ and $y \in Y$, find $x \in X$ for which $f(x) = y$, i.e., solve the equation

$$f(?) = y. \tag{1.1}$$

Existence occurs if this equation has a solution for every $y \in Y$, i.e., if $\operatorname{ran} f = \operatorname{tar} f$. **Uniqueness** occurs if there is at most one solution for every $y \in Y$, i.e., if $f(x) = f(z)$ implies that $x = z$, i.e., the assignment $(f(x) : x \in X)$ is 1-1.

Here are the corresponding map properties:

Definition: The map $f : X \to Y$ is **onto** in case $\operatorname{ran} f = Y$.

Definition: The map $f : X \to Y$ is **1-1** in case $f(x) = f(y) \Longrightarrow x = y$.

Not surprisingly, these two map properties will play a major role throughout these notes. (At last count, '1-1' appears over 300 times in these notes, and 'onto' over 200 times.) − There are other names in use for these properties: An onto map is also called **surjective** or **epimorph(ic)**, while a 1-1 map is also called **injective** or **monomorph(ic)**.

Perhaps the simplest useful examples of maps are those derived from lists, i.e., maps from some $\underline{n}$ into some set Y. Here is the basic observation concerning such maps being 1-1 or onto.

(1.2) If $g : \underline{n} \to Y$ is 1-1 and $f : \underline{m} \to Y$ is onto, then $n \le m$, with equality if and only if g is also onto and f is also 1-1.

Proof: The sequence $(f(1), \ldots, f(m))$ contains every element of Y, but may also contain duplicates of some. Throw out all duplicates to arrive at the sequence $(h(1), \ldots, h(q))$ which still contains all elements of Y but each one only once. In effect, we have 'thinned' f to a map $h : \underline{q} \to Y$ that is still onto but also 1-1. In particular, $q \le m$, with equality if and only if there were no duplicates, i.e., f is also 1-1.

Now remove from $(h(1), \ldots, h(q))$ every entry of the sequence $(g(1), \ldots, g(n))$. Since h is onto and 1-1, each of the n distinct entries $g(j)$ does appear in h's sequence exactly once, hence the remaining sequence $(k(1), \ldots, k(r))$ has length $r = q - n$. Thus, $n \le q$, with equality, i.e., with $r = 0$, if and only if g is onto. In any case, the concatenation $(g(1), \ldots, g(n), k(1), \ldots, k(r))$ provides an 'extension' of the 1-1 map g to a map to Y that is still 1-1 but also onto.

Put the two arguments together to get that $n \le q \le m$, with equality if and only if f is also 1-1 and g is also onto. □

Note the particular conclusion that if both $g : \underline{n} \to Y$ and $f : \underline{m} \to Y$ are 1-1 and onto, then necessarily $n = m$. This number is called the **cardinality** of Y and is denoted

$$\#Y.$$

Hence, if we know that $\#Y = n$, i.e., that there is some invertible map from $\underline{n}$ to Y, then we know that any map $f : \underline{n} \to Y$ is onto if and only if it is 1-1. This is the

(1.3) Pigeonhole principle: If $f : \underline{n} \to Y$ with $\#Y = n$, then f is 1-1 if and only if f is onto.

Any map from $\underline{n}$ to $\underline{n}$ that is 1-1 and onto is called a **permutation of order** n since its list is a reordering of the first n integers. Thus $(3, 2, 1)$ or $(3, 1, 2)$ are permutations of order 3 while the map into $\underline{3}$ given by the 3-vector $(3, 3, 1)$ is not a permutation, as it is neither 1-1 nor onto.

By the Pigeonhole principle, in order to check whether an n-vector represents a permutation, we only have to check whether its range is $\underline{n}$ (which would mean that it is onto, as a map into $\underline{n}$), or we only have to check whether all its values are different and in $\underline{n}$ (which would mean that it is a 1-1 map into its domain, $\underline{n}$).

The finiteness of $\underline{n}$ is essential here. For example, consider the **right shift**

$$r : \mathbb{N} \to \mathbb{N} : n \mapsto n + 1. \tag{1.4}$$

This maps different numbers to different numbers, i.e., is 1-1, but fails to be onto since the number 1 is not in its range. On the other hand, the **left shift**

$$l : \mathbb{N} \to \mathbb{N} : n \mapsto \max\{n - 1, 1\} \tag{1.5}$$

is onto, but fails to be 1-1 since it maps both 1 and 2 to 1.

In light of this example, it is all the more impressive that such a pigeonhole principle continues to hold for certain special maps $f : X \to Y$ with both X and Y infinite. Specifically, according to (4.16)Corollary, if X and Y are *vector spaces* of the same finite *dimension* and $f : X \to Y$ is a *linear* map, then f is 1-1 if and only f is onto. This result is one of the high points of basic linear algebra. A more down-to-earth formulation of it, as in (3.16)Theorem, is the following: *A linear system with as many equations as unknowns has a solution for every right-hand side if and only if it has only the trivial solution when the right-hand side is 0.*

1.13 Prove: any $g : \underline{n} \to Y$ with $n > \#Y$ cannot be 1-1.

1.14 Prove: any $f : \underline{m} \to Y$ with $m < \#Y$ cannot be onto.

1.15 Let $g : \underline{n} \to Y$ be 1-1, and $f : \underline{m} \to Y$ be onto. Prove that

(i) for some $k \geq n$, g can be 'extended' to a map $h : \underline{k} \to Y$ that is 1-1 and onto;

(ii) for some $k \leq m$, f can be 'thinned' to a map $h : \underline{k} \to Y$ that is onto and 1-1.

1.16 Prove: *If T is finite and $S \subset T$, then S is finite, too.* (Hint: consider the set N of all $n \in \mathbb{N} \cup \{0\}$ for which there is a 1-1 map $g : \underline{n} \to S$.)

1.17 Prove that $S \subset T$ and $\#T < \infty$ implies that $\#S \leq \#T$, with equality if and only if $S = T$.

Some examples

The next simplest maps after those given by lists are probably those that come to you in the form of a *list of pairs*. For example, at the end of the semester, I am forced to make up a grade map. The authorities send me the domain of that map, namely the students in this class, in the form of a list, and ask me to assign, to each student, a grade, thus making up a list of pairs of the form

name | grade

Here at UW, the target of the grade map is the set

$$\{\text{A, AB, B, BC, C, D, F, I}\},$$

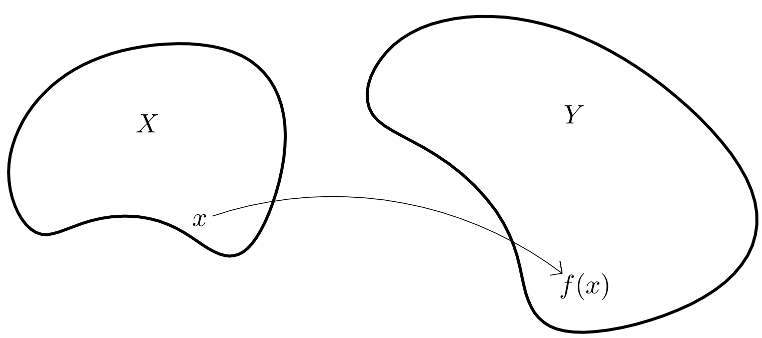
but there is no requirement to make this map onto. In fact, I could not meet that requirement if there were fewer than 8 students in the class. Neither is it required to make the grade map 1-1. In fact, it is not possible to make the grade map 1-1 if the class has more than 8 students in it. But if the class has exactly 8 students in it, then a grade map that is onto is automatically also 1-1, and a grade map that is 1-1 is automatically also onto.

There are many maps in your life that are given as a list of pairs, such as the list of dorm-room assignments or the price list in the cafeteria. The dorm-room assignment list usually has the set of students wanting a dorm room as its domain and the set of available dorm rooms as its target, is typically not 1-1, but the authorities would like it to be onto. The price list at the cafeteria has all the items for sale as its domain, and the set $\mathbb{N}/100 := \{m/100 : m \in \mathbb{N}\}$ of all positive reals with at most two digits after the decimal point as its target. There is little sense in wondering whether this map is 1-1 or onto.

1.18 Describe an interesting map (not already discussed in class) that you have made use of in the last month or so (or, if nothing comes to mind, a map that someone like you might have used recently). Be sure to include domain and target of your map in your description and state whether or not it is 1-1, onto.

Maps and their graphs

One successful mental image of a 'map' is to imagine both domain and target as sets of some possibly indistinct shape, with curved arrows indicating with which particular element in the target the map f associates a particular element in the domain.



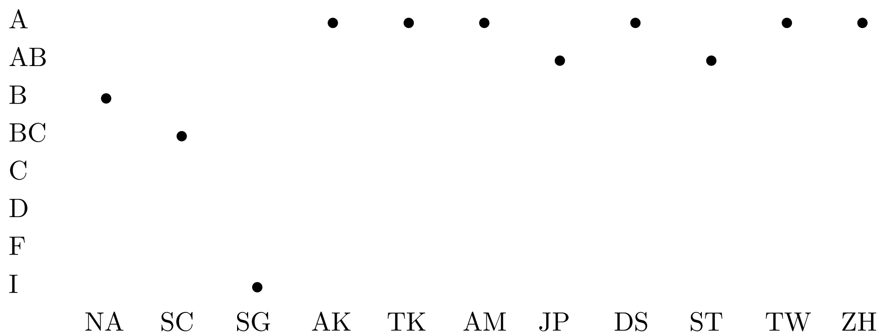
One way to visualize the map $f : X \to Y : x \mapsto f(x)$.

Another successful mental (and more successful mathematical) image of a map $f : X \to Y$ is in terms of its **graph**, i.e., in terms of the set of pairs

$$\{(x, f(x)) : x \in X\}.$$

In fact, the mathematically most satisfying definition of 'map from X to Y' is: *a subset of $X \times Y$ that, for each $x \in X$, contains exactly one pair (x, y)*. In this view, *a map is its graph*.

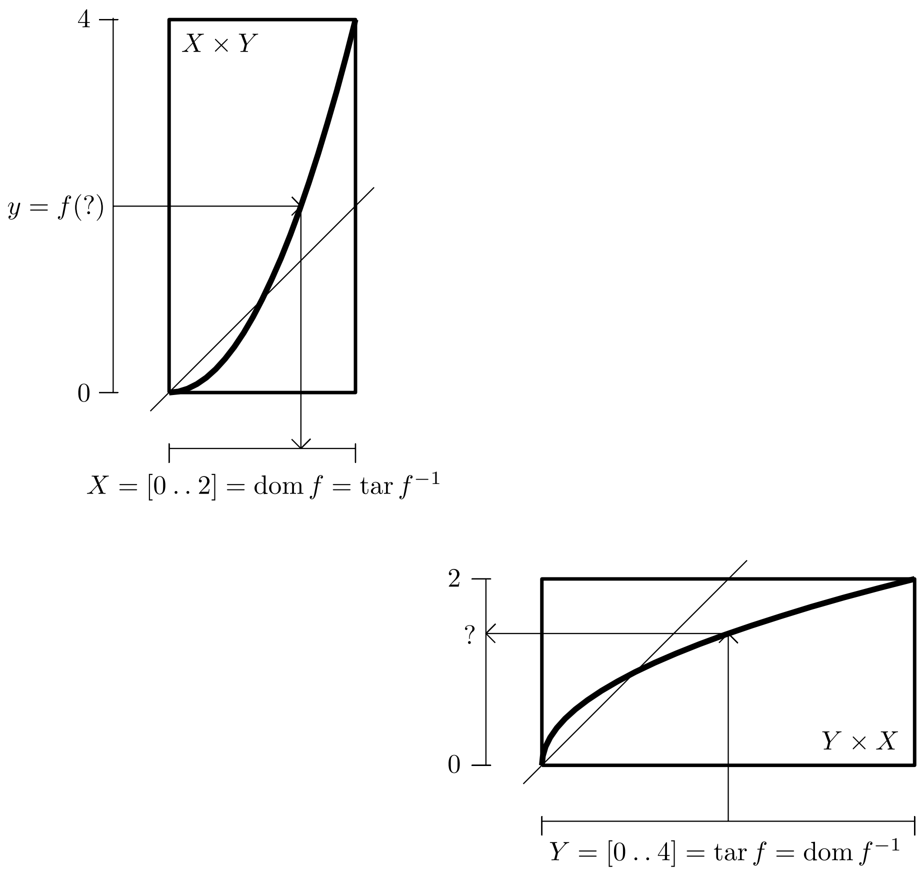
Here, for example, is the (graph of the) grade map G for a graduate course I taught recently. I abbreviated the students' names, to protect the innocent.



You may be more familiar with the graphs of real functions, such as the 'squaring' map

$$()^2 : [0 \,..\, 2] \to [0 \,..\, 4] : x \mapsto x^2,$$

whose graph is shown in the next figure.



The graph of the squaring map $f := ()^2 : [0 \,..\, 2] \to [0 \,..\, 4] : x \mapsto x^2$ and of its inverse $f^{-1} = \sqrt{\ } : [0 \,..\, 4] \to [0 \,..\, 2] : x \mapsto \sqrt{x}$.

1.19 For each of the following subsets R of the cartesian product $X \times Y$ with $X = [0\,..\,2]$ and $Y = [0\,..\,4]$, determine whether it is the graph of a map from X to Y and, if it is, whether that map is 1-1 and/or onto or neither.

(a) $R = \{(x, y) : y = (x - 1/2)^2\}$; (b) $R = \{(x, y) : x \ge 1, y = (2x - 2)^2\}$; (c) $R = \{(x, y) : y = (2x - 2)^2\}$; (d) $R = \{(x, y) : x = y\}$.

1.20 Same as previous problem, but with X and Y interchanged and, correspondingly, R replaced by $R^{-1} := \{(y, x) \in Y \times X : (x, y) \in R\}$. Also, discuss any connections you see between the answers in these two problems.

Invertibility

The graph of a map f helps us solve the standard 'computational' problem involving maps, namely the problem of finding an $x \in X$ that solves the equation

$$f(?) = y$$

for given $f : X \to Y$ and $y \in Y$. The solution set is the pre-image of $\{y\}$ under f, i.e., the set

$$f^{-1}\{y\} = \{x \in X : f(x) = y\}.$$

For example, when looking at the graph of the above grade map G, we see that $G^{-1}\{\text{AB}\} = \{\text{JP, ST}\}$, while $G^{-1}\{\text{D}\} = \{\}$ (the empty set). In the first case, we have two solutions, in the second case, we have none.

In effect, when looking for solutions to the equation $f(?) = y$, we are looking at the graph of f with the roles of domain and target interchanged: We are trying to associate with each $y \in Y$ some $x \in X$ in such a way that $f(x) = y$. If f is onto, then there is *at least* one solution for every $y \in Y$, and conversely (**existence**). If f is 1-1, then there is *at most* one solution for any $y \in Y$, and conversely (**uniqueness**). Ideally, there is, for each $y \in Y$, exactly one $x \in X$ for which $f(x) = y$.

Definition: The map $f : X \to Y$ is **invertible** := for every $y \in Y$ there exists exactly one $x \in X$ for which $f(x) = y$.

Let $f : X \to Y$.

f is invertible if and only if f is 1-1 and onto.

f is invertible if and only if the **inverse of its graph**, i.e., the set

$$\{(f(x), x) : x \in X\} \subset Y \times X,$$

is the graph of a map from Y to X. This latter map is called the **inverse of** f and is denoted by f^{-1}.

Any 1-1 assignment f, taken as a map into its range, is invertible, since it is both 1-1 and onto. The above grade map G fails on both counts to be invertible, it is neither 1-1 nor onto. The squaring map $()^2 : [0\,..\,2] \to [0\,..\,4] : x \mapsto x^2$, on the other hand, is invertible since it is both 1-1 and onto. The earlier figure shows the graph of its inverse, obtained from the graph of the squaring map by reversing the roles of domain and target. In effect, we obtain the inverse of the graph of f by looking at the graph of f sideways and can often tell at a glance whether or not it is the graph of a map, i.e., whether f is 1-1 and onto.

A map may be 'half' invertible, i.e., it may be either 1-1 or onto, without being both. For example, the right shift (1.4) is 1-1, but not onto, while the left shift (1.5) is onto, but not 1-1. Only if domain and target happen to have the same *finite* number of elements, then being 1-1 is guaranteed to be the same as being onto, by the pigeonhole principle (see Problem 1.33).

(1.6) If $f : X \to Y$, with $\#X = \#Y < \infty$, then f 1-1 *or* onto implies f 1-1 *and* onto, i.e., invertible.

In particular, for any *finite* X, any map $f : X \to X$ that is 1-1 *or* onto is automatically invertible.

The notion of f being 'half' invertible is made precise by the notions of left and right inverse. Their definition requires the **identity map**, often written

$$\mathrm{id}$$

if its domain (which is also its target) is clear from the context. The full definition is:

$$\mathrm{id}_X : X \to X : x \mapsto x.$$

In other words, the identity map is a particularly boring map, it leaves everything unchanged.

We also need **map composition**:

Definition: The **composition** $f \circ g$ of two maps $f : X \to Y$ and $g : U \to W \subset X$ is the map

$$f \circ g : U \to Y : u \mapsto f(g(u)).$$

We write fg instead of $f \circ g$ whenever there is no danger of confusion.
Map composition is **associative**, i.e., whenever fg and gh are defined, then

$$(fg)h = f \circ (gh).$$

There is a corresponding definition for the composition $x \circ y$ of two assignments, x and y, under the assumption that $\mathrm{ran}\, y \subset \mathrm{dom}\, x$. Thus,

$$x_y := x \circ y = (x_{y_i} : i \in \mathrm{dom}\, y)$$

is an assignment whose domain is $\mathrm{dom}\, y$ and whose range is contained in $\mathrm{ran}\, x$.

As a simple *example*, if x is an n-vector and y is an m-vector with $\mathrm{ran}\, y \subset \underline{n} = \{1, \ldots, n\}$, then

$$z := x_y := x \circ y = (x_{y_1}, \ldots, x_{y_m}).$$

In `MATLAB`, if `x` describes the n-vector x and `y` describes the m-vector y with entries in $\underline{n} = \{1, \ldots, n\}$, then `z=x(y)` describes the m-vector $z = x_y = x \circ y$.

In the same way, if $\mathtt{A} \in \mathbb{F}^{m \times n}$, and `b` is a k-list with entries from $\underline{m} = \{1, \ldots, m\}$, and `c` is an l-list with entries from $\underline{n} = \{1, \ldots, n\}$, then $\mathtt{A}(\mathtt{b}, \mathtt{c})$ is a $k \times l$-matrix, namely the matrix $\mathtt{D} := \mathtt{A}(\mathtt{b}, \mathtt{c}) \in \mathbb{F}^{k \times l}$ with

$$\mathtt{D}(i, j) = \mathtt{A}(\mathtt{b}(i), \mathtt{c}(j)), \quad i \in \underline{k}, j \in \underline{l}.$$

In effect, the matrix $\mathtt{D} = \mathtt{A}(\mathtt{b}, \mathtt{c})$ is obtained from `A` by choosing rows $\mathtt{b}(1), \mathtt{b}(2), \ldots, \mathtt{b}(k)$ and columns $\mathtt{c}(1), \mathtt{c}(2), \ldots, \mathtt{c}(l)$ of `A`, in that order.

If *all* rows, in their natural order, are to be chosen, then use `A(:,c)`. Again if *all* columns, in their natural order, are to be chosen, then use `A(b,:)`.

In particular, `A(1,:)` is the matrix having the first row of `A` as its sole row, and `A(:,end)` is the matrix having the last column of `A` as its sole column. The matrix `A(1:2:end,:)` is made up from all the odd rows of `A`. `A(end:-1:1,:)` is the matrix obtained from `A` by reversing the order of the rows (as could also be obtained by the command `fliplr(A)`). `A(:,2:2:end)` is obtained by removing from `A` all odd-numbered columns. If `x` is a one-row matrix, then `x(ones(1,m),:)` and `x(ones(m,1),:)` both give the matrix having all its `m` rows equal to the single row in `x` (as would the expression `repmat(x,m,1)`).

`MATLAB` permits the expression `A(b,c)` to appear on the *left* of the equality sign: If `A(b,c)` and `D` are matrices of the same size, then the statement

```
A(b,c) = D;
```

changes, for each (`i`,`j`) $\in$ dom`D`, the entry `A(b(i),c(j))` of `A` to the value of `D(i,j)`. What if, e.g., `b` is not 1-1? `MATLAB` does the replacement for each entry of `b`, from the first to the last. Hence, the last time is the one that sticks. For example, if `a=1:4`, then the statement `a([2 2 2])=[1 2 3]` changes `a` to `[1 3 3 4]`. On the other hand, if `A` appears on both sides of such an assignment, then the one on the right is taken to be as it is at the outset of that assignment. For example,

```
A([i,j],:) = A([j,i],:);
```

is a nice way to interchange the ith row of `A` with its jth.

□

As a first use of map composition, here is the standard test for a map being onto or 1-1.

If fg is onto, then f is onto; if fg is 1-1, then g is 1-1.

Proof: Since $\operatorname{ran}(fg) \subset \operatorname{ran} f \subset \operatorname{tar} f = \operatorname{tar} fg$, fg onto implies f onto. Also, if $g(y) = g(z)$, then $(fg)(y) = (fg)(z)$, hence fg 1-1 implies $y = z$, i.e., g is 1-1. □

For example, the composition lr of the left shift (1.5) with the right shift (1.4) is the identity, hence l is onto and r is 1-1 (as observed earlier).

Remark. The only practical way to check whether a given g is 1-1 is to come up with an f so that fg is 'obviously' 1-1, e.g., invertible. The only practical way to check whether a given f is onto is to come up with a g so that fg is 'obviously' onto, e.g., invertible.

Definition: If $f \in Y^X$ and $g \in X^Y$ and $fg = \mathrm{id}$, then f (being to the left of g) is a **left inverse** of g, and g is a **right inverse** of f. In particular, any left inverse is onto and any right inverse is 1-1.

To help you remember which of f and g is onto and which is 1-1 in case $fg = \mathrm{id}$, keep in mind that being onto provides conclusions about elements of the target of the map while being 1-1 provides conclusions about elements in the domain of the map.

Now we consider the converse statements.

If $f : X \to Y$ is 1-1, then f has a left inverse.

Proof: If f is 1-1 and $x \in X$ is some element, then

$$g : Y \to X : y \mapsto \begin{cases} f^{-1}\{y\} & \text{if } y \in \operatorname{ran} f; \\ x & \text{otherwise,} \end{cases}$$

is well-defined since each $y \in \operatorname{ran} f$ is the image of exactly one element of X. With g so defined, $gf = \mathrm{id}$ follows. □

The corresponding statement: *If $f : X \to Y$ is onto, then f has a right inverse* would have the following 'proof': Since f is onto, we can define $g : Y \to X : y \mapsto$ some point in $f^{-1}\{y\}$. Regardless of how we pick that point $g(y) \in f^{-1}\{y\}$, the resulting map is a right inverse for f. – Some object to this argument since it requires us to pick, for each y, a particular element from that set $f^{-1}\{y\}$. The **belief** that this can *always* be done is known as "The Axiom of Choice".

> If f is an invertible map, then f^{-1} is both a right inverse and a left inverse for f. Conversely, if g is a right inverse for f and h is a left inverse for f, then f is invertible and $h = f^{-1} = g$.
> Consequently, if f is invertible, then: (i)f^{-1} is also invertible, and $(f^{-1})^{-1} = f$; and, (ii) if also g is an invertible map, with $\operatorname{tar} g = \operatorname{dom} f$, then fg is invertible, and $(fg)^{-1} = g^{-1}f^{-1}$ (note the order reversal).

Proof: Let $f : X \to Y$ be invertible. Since, for every $y \in Y$, $f^{-1}(y)$ solves the equation $f(?) = y$, we have $ff^{-1} = \mathrm{id}_Y$, while, for any $x \in X$, x is a solution of the equation $f(?) = f(x)$, hence necessarily $x = f^{-1}(f(x))$, thus also $f^{-1}f = \mathrm{id}_X$.

As to the converse, if f and has both a left and a right inverse, then it must be both 1-1 and onto, hence invertible. Further, if $hf = \mathrm{id}_X$ and $fg = \mathrm{id}_Y$, then (using the associativity of map composition),

$$h = h\,\mathrm{id}_Y = h \circ (fg) = (hf)g = \mathrm{id}_X g = g,$$

showing that $h = g$, hence $h = f^{-1} = g$.

As to the consequences, the identities $ff^{-1} = \mathrm{id}_Y$ and $f^{-1}f = \mathrm{id}_X$ explicitly identify f as a right and left inverse for f^{-1}, hence f must be the inverse of f^{-1}. Also, by map associativity, $(fg)g^{-1}f^{-1} = f\,\mathrm{id}_X f^{-1} = ff^{-1} = \mathrm{id}_Y$, etc. □

While $fg = \mathrm{id}$ implies $gf = \mathrm{id}$ in general only in case $\#\operatorname{dom} f = \#\operatorname{tar} f < \infty$, it does imply that gf is as much of an identity map as it can be: Indeed, if $fg = \mathrm{id}$, then $(gf)g = g \circ (fg) = g\,\mathrm{id} = g$, showing that $(gf)x = x$ for every $x \in \operatorname{ran} g$. There is no such hope for $x \notin \operatorname{ran} g$, since such x cannot possibly be in $\operatorname{ran} gf = g(\operatorname{ran} f) \subset \operatorname{ran} g$. However, since $gf(x) = x$ for all $x \in \operatorname{ran} g$, we conclude that $\operatorname{ran} gf = \operatorname{ran} g$. This makes gf the identity on its range, $\operatorname{ran} g$. In particular, $(gf) \circ (gf) = gf$, i.e., gf is **idempotent** or, a **projector**.

> **(1.7) Proposition:** If $f : X \to Y$ and $fg = \mathrm{id}_Y$, then gf is a projector, i.e., the identity on its range, and that range equals $\operatorname{ran} g$.

For example, the composition lr of the left shift (1.5) with the right shift (1.4) is the identity, hence rl must be the identity on $\operatorname{ran} r = \{2, 3, \ldots\}$ and, indeed, it is.

If the n-vector `c` in `MATLAB` describes a permutation, i.e., if the map $c : \underline{n} \to \underline{n} : j \mapsto$ `c(`j`)` is 1-1 or onto, hence invertible, then the n-vector `cinv` giving its inverse can be obtained with the commands

```
cinv = c; cinv(c) = 1:length(c);
```

The first command makes sure that `cinv` starts out as a vector of the same size as `c`. With that, the second command changes `cinv` into one for which `cinv(c) = [1,2,...,length(c)]`. In other words, `cinv` describes a *left* inverse for (the map given by) `c`, hence the inverse (by the pigeonhole principle).

A second, more expensive, way to construct `cinv` is with the help of the command `sort`, as follows:

```
[d, cinv] = sort(c);
```

For (try `help sort`), whether or not `c` describes a permutation, this command produces, in the n-vector `d`, the list of the items in `c` in nondecreasing order, and provides, in `cinv`, the recipe for this re-ordering:

`d(i)=c(cinv(i))`, `i=` $1{:}n$.

In particular, if `c` describes a permutation, then, necessarily, `d` = `[1,2,3,...]`, therefore `c(cinv)` = `[1,2,...,length(c)]`, showing that `cinv` describes a *right* inverse for (the map given by) `c`, hence the inverse (by the pigeonhole principle).

Both of these methods extend, to the construction of a left, respectively a right, inverse, in case the map given by `c` has only a left, respectively a right, inverse.

□

1.21 Let $f : \underline{2} \to \underline{3}$ be given by the list $(2, 3)$, and let $g : \underline{3} \to \underline{2}$ be the map given by the list $(2, 1, 2)$.

(a) Describe fg and gf (e.g., by giving their lists).

(b) Verify that fg is a projector, i.e., is the identity on its range.

1.22 For each of the following maps, state whether or not it is 1-1, onto, invertible. Also, describe a right inverse or a left inverse or an inverse for it or else state why such right inverse or left inverse or inverse does not exist.

The maps are specified in various ways, e.g., by giving their list and their target or by giving both domain and target and a rule for constructing their values.

(a) a is the map to $\{1, 2, 3\}$ given by the list $(1, 2, 3)$.

(b) b is the map to $\{1, 2, 3, 4\}$ given by the list $(1, 2, 3)$.

(c) c is the map to $\{1, 2\}$ given by the list $(1, 2, 1)$.

(d) $d : \mathbb{R}^2 \to \mathbb{R} : x \mapsto 2x_1 - 3x_2$.

(e) $f : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (-x_2, x_1)$.

(f) $g : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1 + 2, x_2 - 3)$.

(g) $h : \mathbb{R} \to \mathbb{R}^2 : y \mapsto (y/2, 0)$.

1.23 Verify that, in the preceding problem, $dh =$ id, and explain geometrically why one would call hd a projector.

1.24 Prove: *If $fg = fh$ for $g, h : S \to T$ and with $f : T \to U$ 1-1, then $g = h$.*

1.25 Prove: *If $fh = gh$ for $g, h : T \to U$ and with $h : S \to T$ onto, then $f = g$.*

1.26 If both f and g are maps from $\underline{n}$ to $\underline{n}$, then so are both fg and gf. In particular, for any $f \in \underline{n}^{\underline{n}}$, its power sequence

$$f^0 := \mathrm{id}_{\underline{n}}, f^1 := f, f^2 := f \circ f, f^3 := f \circ f^2, \ldots$$

is well defined. Further, since $\underline{n}^{\underline{n}}$ is finite, the sequence $f^0, f^1, f^2, \ldots$ of powers must eventually repeat itself. In other words, there must be a first r such that $f^r = f^j$ for some $j < r$. Let's call the difference $d := r - j$ between these two exponents the **cycle length** of f.

(a) Find the cycle length for the map given by the sequence $(2, 3, 4, 1, 1)$. (Feel free to use `MATLAB`.)

(b) Also determine the cycle lengths for the following maps:

A:=(2,3,4,5,1);	B:=(2,3,1,5,4);	C:=(1,2,3,4,5);
D:=(2,5,2,2,1);	E:=(2,5,2,5,2);	F:=(2,5,2,2,5).

(c) Given all these examples (and any others you care to try), what is your *guess* as to the special nature of the map f^d in case the cycle length of f is d and f is invertible?

1.27 Finish appropriately the following `MATLAB` function

```
function b = ii(a)
% If  ran(a) = N := {1,2,...,length(a)} , hence  a  describes
% the invertible map
%             f:N --> N : j -> a(j)
% then  b  describes the inverse of  f , i.e., the map  g:N --> N  for which
%  fg = id_N  and  gf = id_N .
% Otherwise, the message
%  The input doesn't describe an invertible map
% is printed and an empty  b  is returned.
```


1.28 Let $f_i : X \to X$ for $i = 1{:}n$, hence $g := f_1 \cdots f_n$ is also a map from X to X. Prove that *g is invertible if and only if each f_i is invertible, and, in that case, $g^{-1} = f_n^{-1} \cdots f_1^{-1}$*. (Note the order reversal!)

1.29 If $f : S \to T$ is invertible, then f has exactly one left inverse. Is the converse true?

1.30 Let g be a left inverse for $f : S \to T$, and assume that $\#S > 1$. Prove that g is the unique left inverse for f iff g is 1-1. (Is the assumption that $\#S > 1$ really needed?)

1.31 Let g be a right inverse for f. Prove that g is the unique right inverse for f iff g is onto.

1.32 If $f : S \to T$ is invertible, then f has exactly one right inverse. Is the converse true?

1.33 Let $n := \#X$, hence there is an invertible $g : \underline{n} \to X$.

(i) Prove: *For any $f : X \to Y$, f is 1-1 (onto) if and only if the map fg is 1-1 (onto).*

(ii) Use (i) to derive (1.6) from (1.3).

The inversion of maps

The notions of 1-1 and onto, and the corresponding notions of right and left inverse, are basic to the discussion of the standard 'computational' problem already mentioned earlier: for $f : X \to Y$ and $y \in Y$, solve

$$f(?) \ = \ y. \tag{1.1}$$

When we try to solve (1.1), we are really trying to find, for each $y \in Y$, some $x \in X$ for which $f(x) = y$, i.e., we are trying to come up with a right inverse for f. *Existence* of a solution for every right side is the same as having f onto, and is ensured by the existence of a right inverse for f. Existence of a left inverse for f ensures *uniqueness*: If $hf = \mathrm{id}$, then $f(x) = f(y)$ implies that $x = h(f(x)) = h(f(y)) = y$. Thus existence of a left inverse implies that f is *1-1*. But existence of a left inverse does *not*, in general, provide a solution.

When f has its domain in $\mathbb{R}^n$ and and its target in $\mathbb{R}^m$, then we can think of solving (1.1) *numerically*. Under the best of circumstances, this still means that we must proceed by *approximation*. The solution is found as the limit of a sequence of solutions to *linear* equations, i.e., equations of the form $A? = b$, with A a *linear* map. This is so because linear (algebraic) equations are the only kind of equations we can actually solve exactly (ignoring roundoff). This is one reason why Linear Algebra is so important. It provides the mathematical structures, namely vector spaces and linear maps, needed to deal efficiently with linear equations and, thereby, with other equations.

1.34 T/F

(a) 0 is a natural number.

(b) $\#\{3,3,3\} = 1$.

(c) $\#(3,3,3) = 3$.

(d) $(\{3,1,3,2,4\} \cap \{3,5,4\}) \cup \{3,3\} = \{4,3,3,3\}$.

(e) If A, B are finite sets, then $\#(A \cup B) = \#A + \#B - \#(A \cap B)$.

(f) $\#\{\} = 1$.

(g) $\{3,3,1,6\}\backslash\{3,1\} = \{3,6\}$.

(h) If $f : X \to X$ for some finite X, then f is 1-1 if and only if f is onto.

(i) The map $f : \underline{3} \to \underline{3}$ given by the list $(3,1,2)$ is invertible, and its inverse is given by the list $(2,3,1)$.

(j) The map $f : \underline{3} \to \underline{2}$ given by the list $(1,2,1)$ has a right inverse.

(k) If $U \subset \operatorname{tar} f$, then f maps $f^{-1}U$ onto U.

(l) The map f is invertible if and only if f^{-1} is the graph of a map.

2. Vector spaces and linear maps

Vector spaces, especially spaces of functions

Linear algebra is concerned with vector spaces. These are sets on which two operations, *vector addition* and *multiplication by a scalar*, are defined in such a way that they satisfy various laws. Here they are, for the record:

(2.1) Definition: To say that X is a **linear space** (of **vector**s), or a **vector space**, over the commutative field $\mathbb{F}$ (of **scalar**s) means that there are two maps, (i) $X \times X \to X : (x, y) \mapsto x + y$ called **(vector) addition**; and (ii) $\mathbb{F} \times X \to X : (\alpha, x) \mapsto \alpha x =: x\alpha$ called **scalar multiplication**, which satisfy the following rules.

(a) X is a **commutative group with respect to addition**; i.e., addition

- (a.1) is **associative**: $x + (y + z) = (x + y) + z$;
- (a.2) is **commutative**: $x + y = y + x$;
- (a.3) has **neutral** element: $\exists 0 \, \forall x \; x + 0 = x$;
- (a.4) has **inverse**: $\forall x \, \exists y \; x + y = 0$.

(s) scalar multiplication is

- (s.1) **associative**: $\alpha(\beta x) = (\alpha\beta)x$;
- (s.2) **field-addition distributive**: $(\alpha + \beta)x = \alpha x + \beta x$;
- (s.3) **vector-addition distributive**: $\alpha(x + y) = \alpha x + \alpha y$;
- (s.4) **unitary**: $1x = x$.

It is standard to denote the element $y \in X$ for which $x + y = 0$ by $-x$ since such y is uniquely determined by the requirement that $x + y = 0$. I will denote the neutral element in X by the same symbol, 0, used for the zero scalar.

While the **scalar**s can come from some abstract field, we will only be interested in the real scalars $\mathbb{R}$ and the complex scalars $\mathbb{C}$. Also, from a practical point of view, the most important linear spaces consist of **function**s, i.e., of scalar-valued maps all on some common domain. This means that the typical linear space we will deal with is (a subset of) the collection of all maps $\mathbb{F}^T$ from some fixed domain T into the **scalar field** $\mathbb{F}$ (either $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$), with **pointwise** addition and multiplication by scalars. Here is the definition: