Chapter 5

Subspaces and Linear Equations

5.1 Subspaces

Lurking inside many of our examples and discussions in the previous chapters, sometimes awkwardly bubbling to the surface unnamed, has been a fundamental concept that now needs to be formulated explicitly.

Definition: A nonempty subset $\mathcal{W}$ of a vector space $\mathcal{V}$ is a *subspace* of $\mathcal{V}$ if $\mathcal{W}$ is *closed* under addition and scalar multiplication:

$$\text{If } \vec{x}, \vec{y} \in \mathcal{W}, \text{ then } \vec{x} + \vec{y} \in \mathcal{W} \text{ and } r\vec{x} \in \mathcal{W} \text{ for all real } r.$$

Theorem 1: A subspace of $\mathcal{V}$ is itself a vector space with the same addition and multiplication operations as $\mathcal{V}$.

This is a corollary of:

Lemma: Every subspace of $\mathcal{V}$ contains the 0 vector of $\mathcal{V}$.

Proof of Lemma: $\vec{x} + (-1)\vec{x} = 0$ for any $\vec{x} \in \mathcal{W}$. Since $\mathcal{W}$, by assumption, is nonempty and is closed under the vector operations, this shows that $0 \in \mathcal{W}$.

All the other parts of the definition of a vector space are identities such as $\vec{x} + \vec{y} = \vec{y} + \vec{x}$. Since they were true before, they are still true now, so the subspace is a vector space. (This proves the theorem.)

Examples

1. Physical 3-dimensional space has the following subspaces:
   a) any line through the origin
   b) any plane through the origin
   c) the set consisting of the origin alone: $\{0\}$
   d) the whole space
2. Subspaces of $\mathbf{R}^3$: These are the same as the things in the previous list, but now regarded numerically.

a) Examples of "lines" (one-dimensional subspaces) are:

[the set of] all vectors of form $(0, y, 0)$;

all vectors of form $(t, 2t, 0) = t(1, 2, 0)$.

b) Examples of "planes" (two-dimensional subspaces) are:

all vectors of form $(x, y, 0)$;

all vectors of form $(2s + 3t, 7t, s - t) = s(2, 0, 1) + t(3, 7, -1)$.

c–d) $\{(0, 0, 0)\}$ and $\mathbf{R}^3$ itself also count as subspaces of $\mathbf{R}^3$.

REMARK: To show that a (nonempty) subset of a vector space is a subspace, *all we need to do is to check that it is closed under the two vector operations*. We don't need to verify the 8 axioms, because we already know they are true (the big space is a vector space!). Let's write out such a check for one of the examples above:

$$\begin{aligned}(t_1, 2t_1, 0) + (t_2, 2t_2, 0) &= (t_1 + t_2, 2(t_1 + t_2), 0) \\ &= (t, 2t, 0) \quad \text{with } t = t_1 + t_2\,;\end{aligned}$$

$$\begin{aligned}r(t, 2t, 0) &= (rt, 2rt, 0) \\ &= (t', 2t', 0) \quad \text{with } t' = rt\,.\end{aligned}$$

Since we have already verified the eight axioms for addition and scalar multiplication of numerical triples, this shows that the set in question is a vector space.

3. $\mathcal{P}_2$ is a subspace of the vector space of all polynomials. The polynomials of the form $at^2 + c$ form a subspace of $\mathcal{P}_2$. Those of the form $at^2 + bt + c$ with $a \neq 0$ *don't*; we saw earlier (Sec. 3.1) that this set is not closed under addition.

4. An example of importance in electrostatics (in solving Laplace's equation, $\nabla^2\phi = 0$): Consider the space of all polynomials in *two* variables, x and y. Let us call this space $\mathcal{Z}$. (All the vector-space names in this example are temporary conveniences, not standard notations.) A typical element of $\mathcal{Z}$ is $x^4y + y^2x + 3y^2 - 1$. Let $\mathcal{Y}$ be the subspace of $\mathcal{Z}$

consisting of all quadratic polynomials — that is, those of degree up to 2, counting the total contribution of both x and y to a term:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F.$$

(This should stir memories of "conic sections".) Now consider three subspaces of $\mathcal{Y}$:

$\mathcal{Y}_2 \equiv$ the *homogeneous* quadratic polynomials:

$$Ax^2 + Bxy + Cy^2.$$

$\mathcal{Y}_1 \equiv$ the homogeneous linear polynomials: $\quad Dx + Ey.$

$\mathcal{Y}_0 \equiv$ the constant polynomials: $\quad F.$

Notice that:

(1) Every element of $\mathcal{Y}$ is a sum of 3 elements, one from each $\mathcal{Y}_i$ $(i = 0,\ 1,\ 2)$.

(2) Nothing except $\vec{0}$ belongs to more than one $\mathcal{Y}_i$.

In such a situation one says that $\mathcal{Y}$ is the *direct sum* of the subspaces $\mathcal{Y}_i$ and writes

$$\mathcal{Y} = \mathcal{Y}_2 \oplus \mathcal{Y}_1 \oplus \mathcal{Y}_0 \,.$$

Finally, an important subspace of $\mathcal{Y}_2$ is the space $\mathcal{H}_2$ of *harmonic* homogeneous quadratic polynomials — those satisfying the partial differential equation

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} \equiv \nabla^2 p = 0.$$

These are of the form $A(x^2 - y^2) + Bxy$. Similarly, the harmonic homogeneous polynomials of any degree n can be identified; call the space of them $\mathcal{H}_n$. Now, the payoff: It can be shown that *any* solution of Laplace's equation can be approximated by an element of the space

$$\mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \oplus \mathcal{H}_4 \oplus \cdots.$$

In more classical terms, *every solution is a convergent series of harmonic homogeneous polynomials*. (When rewritten in polar coordinates, a function in $\mathcal{H}_n$ will be of the form

$$C_1 r^n \cos n\theta + C_2 r^n \sin n\theta,$$

and it is in this form that students usually first encounter them. The convergent series then becomes a *Fourier series*.)

5. An old example revisited: The solutions of

$$\frac{d^2y}{dt^2} + \omega^2 y = 0$$

(for some nonzero constant ω), which are the functions

$$A\cos\omega t + B\sin\omega t,$$

form a subspace of the space of all continuous functions, $\mathcal{C}(-\infty, \infty)$. A similar statement is true for any *homogeneous linear* differential equation.

REMARK: You may have a feeling that the definition of a subspace is strangely similar to the definition of a linear function. Furthermore, in many examples the calculations that need to be performed to show that a certain set is a subspace may seem to be identical to those that show that a certain function is linear. These similarities cause many students to get the concepts of "linear function" and "subspace" somehow confused. Please keep in mind that a subspace is a set (a collection of vectors) but a function is a mapping from one set into another (taking vectors as input and emitting new vectors as output). In the next section we'll see that if L is a linear function, then the set of vectors satisfying $L\vec{x} = 0$ is always a subspace (of the domain space of L). It turns out that the most obvious subspaces to use as elementary examples in textbooks, homework exercises, and tests are subspaces of precisely this type, and therefore the verification that the set is a subspace is indeed very closely related to the proof that the linear function involved is linear. The difference is that in proving linearity we show that the vectorial *expressions*

$$L(r\vec{x} + \vec{y}) \quad \text{and} \quad rL(\vec{x}) + L(\vec{y})$$

are equivalent, whereas in proving that the solutions of $L\vec{x} = 0$ form a subspace we show that the *equations*

$$L(r\vec{x} + \vec{y}) = 0 \quad \text{and} \quad rL(\vec{x}) + L(\vec{y}) = 0$$

are equivalent. With this in mind, let's look at four more examples.

Example 1: Show that the matrices of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

(where a and b are real numbers) make up a subspace of the vector space of real 2×2 matrices.

SOLUTION: We have already observed in an earlier section that the $n \times m$ matrices do form a vector space. The subset of matrices under consideration is nonempty (for example, it contains the 2×2 zero matrix), so all that must be done is to show that it is closed under the vector operations. Form a linear combination of two arbitrary members of the set:

$$r \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} + \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix},$$

where r stands for an arbitrary real number. By definition of scalar multiplication and addition of matrices, this is

$$\begin{pmatrix} ra_1 + a_2 & -(rb_1 + b_2) \\ rb_1 + b_2 & ra_1 + a_2 \end{pmatrix},$$

which is of the desired general form (with $a = ra_1 + a_2$ and $b = rb_1 + b_2$).

Example 2: Show that the set of all 3×3 matrices X satisfying $XA = 0$, where A is a fixed 3×3 matrix, is a vector space.

SOLUTION: Again, we know that the set of *all* 3×3 matrices is a vector space, so what needs to be proved is just that the set in question is a subspace of that space. And again, the zero matrix satisfies the condition, so the set is not empty. Assume that $X_1A = 0$ and $X_2A = 0$. We need to prove that $(rX_1 + X_2)A = 0$ also. Well, it follows from the algebraic properties of matrix multiplication that

$$(rX_1 + X_2)A = r\,X_1A + X_2A, \qquad (*)$$

and this is 0 by the assumptions. Note that the calculation $(*)$ is nothing but the proof that the function

$$L(X) \equiv XA$$

is linear. (Note also that writing out matrices as actual 3×3 arrays of numbers or variables would be more of a hindrance than a help in discovering and (especially) explaining this proof.)

Example 3: Show that $\mathcal{C}^2(a,b)$ is a subspace of $\mathcal{C}(a,b)$. (Here (a,b) is some interval (possibly infinite) in the real line, $\mathcal{C}(a,b)$ is the space of all continuous functions with that interval as domain, and $\mathcal{C}^2(a,b)$ is the set of all functions on that domain whose second derivatives exist and are continuous.)

SOLUTION: If a function is differentiable, then it certainly is continuous. (This is a familiar theorem of calculus.) This shows that $\mathcal{C}^2$ is a sub*set* of $\mathcal{C}$ (and we know that twice differentiable functions exist, so the set is not empty). Since the addition and scalar multiplication operations on functions satisfy the eight vectorial conditions, all we need to show is that sums and scalar multiples of twice differentiable functions are twice differentiable. But this follows from standard theorems of calculus (the derivative of a sum is the sum of the derivatives, etc.).

Example 4: Let M be the set of all *even* continuous functions on the domain $[-1,1]$. Is M a subspace of $\mathcal{C}[-1,1]$?

SOLUTION: "Even" means that $p(-t) = p(t)$. Let p and q be any even functions from $\mathcal{C}[-1,1]$. Then we will have

$$(rp+q)(-t) = rp(-t) + q(-t) = rp(t) + q(t) = (rp+q)(t).$$

So the set is closed under the vector operations.

Example 5: Show that the solutions of the differential equation $y'' = 9y$ are a subspace of $\mathcal{C}^2(-\infty,\infty)$.

SOLUTION: Obviously any solution must be twice differentiable, so the solutions are a subset of $\mathcal{C}^2$. If y_1 and y_2 are solutions, then so is $ry_1 + y_2$, as we can quickly verify:

$$(ry_1 + y_2)'' = ry_1'' + y_2'' = r(9y_1) + 9y_2 = 9(ry_1 + y_2)\,.$$

Equivalently, we rewrite the equation as

$$L(y) \equiv y'' - 9y = 0$$

and verify that L is linear (by essentially the same calculation as just given).

A more concrete proof is also possible, of course: We know that the solutions are the linear combinations of the functions e^{3t} and e^{-3t}. These functions are twice differentiable. Also, almost by the very definition of

"linear combination", this set of functions is closed under addition and scalar multiplication. This last observation is a special case of our next theorem.

Theorem 2: If $\mathcal{S} = \{\vec{x}_1, \ldots, \vec{x}_n, \ldots\}$ is a subset (finite or infinite) of a vector space $\mathcal{V}$, then the span of $\mathcal{S}$ is a subspace of $\mathcal{V}$. In fact, it is the smallest subspace that contains all of the $\vec{x}_j$ themselves.

We leave the proof of this theorem and the next (and an explanation of what's meant by "smallest") for the exercises. (See Sec. 4.1 for the definition of a span.)

Theorem 3: If $\mathcal{S}$ is a subspace already, then $\operatorname{span}\mathcal{S} = \mathcal{S}$; and conversely.

Theorem 4: If $\mathcal{U}$ is a subspace of $\mathcal{V}$, then $\dim\mathcal{U} \leq \dim\mathcal{V}$. In fact, $\dim\mathcal{U} < \dim\mathcal{V}$, unless either $\mathcal{U} = \mathcal{V}$ or $\dim\mathcal{U} = \infty$.

PROOF: Use the theorems in Sec. 4.3: If $\mathcal{U}$ is finite-dimensional, then $\mathcal{U}$ has a basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, where $k = \dim\mathcal{U}$. Since $\mathcal{B}$ is a basis, it is linearly independent. Similarly, if $\mathcal{U}$ is infinite-dimensional, then it includes an infinite independent set $\mathcal{B}$. By adding vectors to $\mathcal{B}$, one can construct a basis for $\mathcal{V}$. Obviously, the number of vectors in this basis is greater than $\dim\mathcal{U}$.

AFFINE SUBSPACES

Definition: If $\mathcal{W}$ is a subspace of $\mathcal{V}$ and $\vec{p}$ is a vector in $\mathcal{V}$, then the set $\mathcal{A}$ of all vectors of the form $\vec{x} + \vec{p}$ with $\vec{x} \in \mathcal{W}$ is an *affine subspace* of $\mathcal{V}$.

If $\vec{p} \in \mathcal{W}$, then $\mathcal{A} = \mathcal{W}$. If $\vec{p} \notin \mathcal{W}$, then $\mathcal{A}$ is *not* a subspace (of $\mathcal{V}$), despite the name (just as a vice president is not a president).

Examples:

1. The affine subspaces of 3-space that are not subspaces are the lines and planes that do *not* pass through the origin, and the points that are not the origin. (Compare the parametric expression for a plane (Sec. 1.2), $\vec{x} = \vec{v} + s\vec{u}_1 + t\vec{u}_2$. Here $\mathcal{W} = \operatorname{span}\{\vec{u}_1, \vec{u}_2\}$ and $\vec{p} = \vec{v}$.) The fact that these are geometrically equivalent to lines, planes, and points that *do* pass through the origin is a reflection of the arbitrariness of the location of the origin. (That is, physical space is not a vector space *a priori*; it becomes one only when an origin of coordinates is chosen. Cf. Sec. 1.3.)

2. The space of all solutions of the nonhomogeneous linear differential equation
$$\frac{d^2y}{dt^2} + \omega^2 y = -\cos 2\omega t$$
is an affine subspace of $\mathcal{C}(-\infty, \infty)$. They are the solutions of the form
$$\begin{aligned} y &= A\cos\omega t + B\sin\omega t + \frac{1}{3\omega^2}\cos 2\omega t \\ &\equiv \vec{x} + \vec{p}, \end{aligned}$$
where $\vec{x}$ is the general solution of the corresponding homogeneous equation (the general element of a true subspace), and $\vec{p}$ is a particular solution of the nonhomogeneous equation. (Note that the $\vec{p}$ in the definition of an affine subspace is itself a member of the affine subspace, since the zero vector is a permissible choice of $\vec{x}$.) In general, the space of solutions of an *nonhomogeneous linear* differential equation is an affine subspace of the vector space of all functions of the proper type to be considered as possible solutions.

Note that both subspaces and affine subspaces are "flat". A more general type of subset of $\mathbf{R}^3$ is a *submanifold* (for instance, a curved surface), but those are beyond the scope of this course.

Intersections, unions, and sums

If $\mathcal{S}$ and $\mathcal{T}$ are any sets, their *intersection*, $\mathcal{S} \cap \mathcal{T}$, is the set of all objects that belong to *both* $\mathcal{S}$ *and* $\mathcal{T}$, and their *union*, $\mathcal{S} \cup \mathcal{T}$, is the set of all objects that belong to *either* $\mathcal{S}$ *or* $\mathcal{T}$. Our interest here is in the situation where the objects in question all belong to some vector space $\mathcal{V}$, and the sets $\mathcal{S}$ and $\mathcal{T}$ either are subspaces of $\mathcal{V}$ or are linearly independent sets that serve as bases for subspaces.

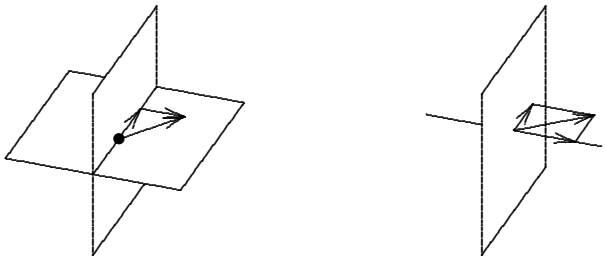
Theorem 5: If $\mathcal{S}$ and $\mathcal{T}$ are subspaces, then so is $\mathcal{S} \cap \mathcal{T}$.

We leave the proof as an exercise. The most important point is that the analogue of this theorem for unions is *false*. Except in the special case where one of the subspaces lies entirely inside the other, the union of two subspaces is not closed under addition. This is obvious from the sketches below, one showing the union of two planes in $\mathbf{R}^3$, the other that of a plane and a line. (The latter sketch shows a vector sum that leads outside the union.)

The subspace naturally associated with the union of two subspaces is the *span* of their union, also called the *sum* of the subspaces:

$$\mathcal{S} + \mathcal{T} \equiv \{\vec{x} \in \mathcal{V} \colon \vec{x} = \vec{s} + \vec{t} \text{ for some } \vec{s} \in \mathcal{S} \text{ and some } \vec{t} \in \mathcal{T}\}.$$

In both cases shown in the two sketches, the sum is the entire vector space $\mathbf{R}^3$. As previously remarked, if the intersection of the subspaces consists of the zero vector alone, then the sum is called *direct* and written $\mathcal{S} \oplus \mathcal{T}$. The uniqueness theorem for the expansion of a vector with respect to a basis has a counterpart for direct sums: Every element of a direct sum can be decomposed in *exactly one way* into a piece from $\mathcal{S}$ and a piece from $\mathcal{T}$. It is clear from the sketches that the sum of the line and the plane is direct, but the sum of the two planes is not. (The vectors in the left sketch show the nonuniqueness of the decomposition.)



The next theorem implies that the sum of two planes in 3-dimensional space is never direct, but that the sum of a plane and a line is always direct unless the line lies in the plane. (Since we are discussing vector subspaces, we must consider only lines and planes that pass through the origin.)

Theorem 6: If $\mathcal{S}$ and $\mathcal{T}$ are subspaces of a finite-dimensional vector space $\mathcal{V}$, then

$$\dim(\mathcal{S} + \mathcal{T}) = \dim \mathcal{S} + \dim \mathcal{T} - \dim(\mathcal{S} \cap \mathcal{T}).$$

In particular, if the sum is direct, one has

$$\dim(\mathcal{S} \oplus \mathcal{T}) = \dim \mathcal{S} + \dim \mathcal{T}.$$

REMARK: To gain some appreciation for the proof of this theorem, let's first contemplate the problem of finding bases for $\mathcal{S} \cap \mathcal{T}$ and $\mathcal{S} + \mathcal{T}$ if we are given bases for $\mathcal{S}$ and $\mathcal{T}$. Assume that

$$\mathcal{S} = \text{span}\{s_1, \ldots, s_k\} \equiv \text{span}\, S \quad \text{and} \quad \mathcal{T} = \text{span}\{t_1, \ldots, t_l\} \equiv \text{span}\, T$$

and that $k = \dim \mathcal{S}$, $l = \dim \mathcal{T}$ (which is equivalent to the condition that each list, S and T, be linearly independent). To get a basis for $\mathcal{S} \cap \mathcal{T}$, one's first thought might be to look at the intersection of the bases, $S \cap T$. Unfortunately, that set is likely to be empty, even if $\mathcal{S} \cap \mathcal{T}$ is nontrivial. For example, consider the intersecting* planes

$$\mathcal{S} = \text{span}\{(1,2,3), (1,0,0)\}, \qquad \mathcal{T} = \text{span}\{(0,1,0), (2,2,2)\}.$$

The situation with respect to the sum is slightly better, but not ideal: The union of the bases, $S \cup T$, is certainly a spanning set for $\mathcal{S} + \mathcal{T}$, but it is not necessarily an independent set. In this case we have an algorithm from Chapter 4 for replacing it by an independent set, which will be the desired basis for $\mathcal{S} + \mathcal{T}$.

PROOF: The trick is to be more careful in our choice of bases. Let $U \equiv \{\vec{u}_1, \ldots, \vec{u}_j\}$ be a basis for $\mathcal{S} \cap \mathcal{T}$. Then (see below) we can extend U to get a basis for $\mathcal{S}$,

$$S \equiv \{\vec{u}_1, \ldots, \vec{u}_j, \vec{s}_{j+1}, \ldots, \vec{s}_k\}$$

(in other words, take $\vec{s}_i \equiv \vec{u}_i$ for $i \leq j$). Similarly, we can extend U to get a basis for $\mathcal{T}$,

$$T \equiv \{\vec{u}_1, \ldots, \vec{u}_j, \vec{t}_{j+1}, \ldots, \vec{t}_l\}.$$

By definition of "dimension", the numbers that arose here are

$$j = \dim(\mathcal{S} \cap \mathcal{T}), \quad k = \dim \mathcal{S}, \quad l = \dim \mathcal{T}.$$

Now throw together all three categories of basis vectors ($\vec{u}$s, $\vec{s}$s, and $\vec{t}$s) into one big list, $S \cup T$. (Do not count the $\vec{u}$s twice!) The number of vectors in $S \cup T$ is $k + l - j$ (right?). Therefore, the theorem will be proved if we can

* The normal vectors to $\mathcal{S}$ and $\mathcal{T}$, computed from the cross product, are $(0,3,-2)$ and $(2,0,-2)$ respectively. Since these are not parallel, the two planes intersect in a line (rather than coinciding).

establish that $S \cup T$ is a basis for $\mathcal{S} + \mathcal{T}$. It is obviously a spanning set (see remark above), so we just need to check that it is independent. Suppose that[†]

$$\begin{aligned} 0 &= \sum_{i=1}^{j} c_i \vec{u}_i + \sum_{i=j+1}^{k} c_i \vec{s}_i + \sum_{i=k+1}^{k+l-j} c_i \vec{t}_{i+j-k} \\ &\equiv \vec{u} + \vec{s} + \vec{t}. \end{aligned}$$

Then

$$-\vec{t} = \vec{u} + \vec{s},$$

and this is a vector that lies in both $\mathcal{T}$ and $\mathcal{S}$ (because $\vec{t} \in \mathcal{T}$, $\vec{s} \in \mathcal{S}$, $\vec{u} \in \mathcal{S} \cap \mathcal{T} \subseteq \mathcal{S}$). Therefore, $-\vec{t}$ belongs to $\mathcal{S} \cap \mathcal{T}$. Therefore, $-\vec{t}$ is a linear combination of the vectors $\vec{u}_i$. But since the expansion of $-\vec{t}$ (as an element of $\mathcal{S}$) as a linear combination of the basis S is unique, this means that the coefficients c_i for $j < i \leq k$ (i.e., the vector $\vec{s}$) must actually be zero. In exactly the same way we could have shown that $\vec{t}$ is zero. Therefore, $\vec{u} = -(\vec{s} + \vec{t})$ is zero. That is, *all* the coefficients c_i are zero, which proves independence of the basis and finishes the proof of the theorem.

In proving Theorems 4 and 6 we have used a fairly obvious fact that is important enough to be stated separately:

Lemma: If $\mathcal{S}$ is a subspace of $\mathcal{V}$, then any basis for $\mathcal{S}$ can be made into a basis for $\mathcal{V}$ by adding vectors.

PROOF: This is an immediate corollary of Proposition 2(b) in Sec. 4.3.

Exercises

5.1.1 Show that the span of a set of vectors is a subspace. (You may assume that the original set is not empty.)

5.1.2 Justify the statement that the span of a set of vectors is the *smallest* subspace that contains all of those vectors. (This means that any other subspace that contains all the original vectors also includes the entire span.)

5.1.3 Prove that if $\mathcal{S}$ is a subspace, then its span is itself. Also prove the *converse*: If $\operatorname{span} \mathcal{S} = \mathcal{S}$, then $\mathcal{S}$ is a subspace.

[†] This is merely the most general linear combination of the vectors in $S \cup T$; the notation looks complicated because we had to change the indexing of T to avoid a conflict.

5.1.4 The set of all real 2×2 matrices is a vector space; let's call it $\mathcal{M}$.

(a) What is the dimension of $\mathcal{M}$? (NOTE: This is not the same as the dimension of the space of vectors that the matrices act upon!)

(b) Let $\mathcal{D}$ be the subset of $\mathcal{M}$ consisting of all matrices with determinant 0. Is $\mathcal{D}$ a sub*space* of $\mathcal{M}$?

5.1.5 Let $\mathcal{M}$ be the vector space of 3×3 matrices.

(a) What is the dimension of $\mathcal{M}$?

(b) Show that the set of all symmetric 3×3 matrices (satisfying $A_{kj} = A_{jk}$) is a subspace of $\mathcal{M}$.

(c) What is the dimension of that subspace?

5.1.6 Answer (a) and (c) of the preceding exercise for the case of $n \times n$ matrices.

5.1.7 Prove that the $n \times n$ antisymmetric matrices form a subspace in the space of all square matrices of the same size. Find a basis for this space and the dimension of it.

5.1.8 Show that the solution space of $x + 2y^2 = 0$ is *not* a subspace of $\mathbf{R}^2$.

5.1.9 Which of these subsets of $\mathbf{R}^3$ is a subspace? Explain.

(A) The set of all vectors (x, y, z) such that $x + 2y + 3z = 0$ **and** $x - y = 0$.

(B) The set of all vectors (x, y, z) such that $x + 2y + 3z = 0$ **or** $x - y = 0$.

5.1.10 Let $\mathcal{S}$ be the set of all twice-differentiable functions y on $(-\infty, \infty)$ such that either $y'' + 4y = 0$ or $y'' - 4y = 0$. Show that $\mathcal{S}$ is *not* a subspace of $\mathcal{C}^{(2)}(-\infty, \infty)$.

5.1.11 Which of these is *not* a subspace of the function space $\mathcal{C}(0, 1)$? ($\mathcal{C}(0, 1)$ is the space of continuous real-valued functions on the interval $0 < t < 1$.)

(A) the set of differentiable functions, $\mathcal{C}^1(0, 1)$

(B) the set of (continuous) functions satisfying $f(0.5) = 0$

(C) the solutions of $\frac{d}{dt} f = -5f$ on that interval

(D) the set of (continuous) nonnegative functions (that is, $f(t) \geq 0$ for all t in the interval)

5.1.12 Decide whether each of these is a "true" subspace of $\mathbf{R}^3$, an affine subspace of $\mathbf{R}^3$, or neither.

(a) the solutions of $5x + 5y - 2z = 10$

(b) the vectors (x, y, z) with $y = 3$

(c) the plane defined by $x = -y$, with z arbitrary

(d) the solutions of $x + y^3 = -\sin(2\pi z)$

5.1.13 $\mathcal{C} \equiv \mathcal{C}^0(-\infty, \infty)$ is the space of continuous, real-valued functions of a real variable. Tell which of the following subsets of $\mathcal{C}$ are sub*spaces*.

(a) the positive functions ($f(t) > 0$ for all t)

(b) the functions satisfying $f(2) = 0$

(c) the solutions of the differential equation $f'(t) + 5f(t) = 0$

(d) the span of the set $\{\sin t, \cos t\}$

5.1.14 Let $\mathcal{N}$ be the set of all odd continuous functions on $[-1, 1]$. Is $\mathcal{N}$ a subspace of $\mathcal{C}[-1, 1]$? (An *odd* function is one satisfying $p(-t) = -p(t)$ for all t in the domain.)

5.1.15 Which of these sets of functions with domain $[-1, 1]$ are subspaces of $\mathcal{C}[-1, 1]$?

(a) the decreasing functions

(b) the functions satisfying the condition $p(0) = 1$

(c) the functions satisfying the condition $\int_{-1}^{1} p(t)\, dt = 0$.

5.1.16 Which of the following sets are real vector spaces? Justify your answer. (When the answer is "yes", the proper justification usually is to show that the set is a subspace of an already familiar vector space.)

(a) the set of all polynomials with rational coefficients

(b) all vectors $(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ whose coordinates satisfy the equation $x_1 + x_2 + \ldots + x_n = 2$

(c) the set of all functions in $\mathcal{C}[0, 3]$ satisfying $p(1) = 0$

(d) all vectors (points) in the first quadrant of the plane $\mathbf{R}^2$.

(e) All vectors $(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ whose coordinates satisfy the equation $x_1 + x_2 + \ldots + x_n = 0$.

5.1.17 Prove Theorem 5 ("The intersection of subspaces is a subspace.").

5.1.18 Consider affine subspaces in $\mathbf{R}^3$ (lines and planes that do not necessarily pass through the origin). Use Theorem 6 to classify the possible intersections and unions of

(a) a plane and a line;

(b) two planes.

HINT: If the sets intersect at all, you can choose the origin of coordinates to be a point in the intersection.

5.1.19 Let $\mathbf{R}^\infty$ denote the space of all *infinite* sequences of real numbers, $\vec{x} = (x_1, x_2, \dots)$. (Any of the x_j are allowed to be zero.) Define the operations of addition and scalar multiplication by

$$(x_1, x_2, \dots, x_n, \dots) + (y_1, y_2, \dots, y_n, \dots) \\ = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n, \dots),$$

$$r(x_1, x_2, \dots, x_n, \dots) = (rx_1, rx_2, \dots, rx_n, \dots).$$

Show that $\mathbf{R}^\infty$ is a vector space (i.e., the 8 axioms are satisfied).

In the remaining exercises, prove that each of these subsets of $\mathbf{R}^\infty$ (defined in the preceding exercise) is a subspace of $\mathbf{R}^\infty$. (See the examples at the end to clarify the definitions.)

5.1.20 $\mathcal{U}_1$ = the set of *bounded* sequences: There is a number M $(< \infty)$ such that $x_n < M$ for all n.

5.1.21 $\mathcal{U}_2$ = the set of sequences that *converge*: $\lim_{n\to\infty} x_n \equiv L$ exists (and is not ∞). HINT: Use the theorem from calculus that

$$\lim_{n\to\infty} (x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n$$

if the two limits on the right exist.

5.1.22 $\mathcal{U}_3$ = the set of sequences that converge to 0.

5.1.23 $\mathcal{U}_4$ = the set of sequences that are *eventually zero*: There is an index N such that $x_n = 0$ for all $n > N$.

EXAMPLES:

$(1, 2, 4, 8, \dots)$ (with the obvious infinite continuation) is an element of $\mathbf{R}^\infty$ that is not in $\mathcal{U}_1$.

$(1, 0, 1, 0, 1, \dots)$ is an element of $\mathcal{U}_1$ that is not in $\mathcal{U}_2$.

$(1.1, 1.01, 1.001, 1.0001, \dots)$ is an element of $\mathcal{U}_2$ that is not in $\mathcal{U}_3$.

$(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots)$ is an element of $\mathcal{U}_3$ that is not in $\mathcal{U}_4$.

$(2, 1, 0, 0, 0, \dots)$ is an element of $\mathcal{U}_4$.

5.2 Subspaces Associated with a Linear Function: Kernel and Range

Earlier (Sec. 3.2) it was pointed out that every linear function is associated with two vector spaces, its domain and codomain. The domain and codomain logically exist prior to the function and form an inert arena where the function acts. More significant are two subspaces of these spaces, the *kernel* and *range*, which are determined by the linear function and give us profound information about the linear function itself. In this section and the next, we not only define these subspaces and develop some related theory, but also stress why they are so important in applications (to differential equations, for example). The main point we wish to emphasize is that these "abstract" concepts and theorems are important practical tools, which *tell you how to solve problems.*

Be warned that the terminology used here, though fairly standard, is not universal. In particular, the word "range" is ambiguous in the mathematical literature. The alternative terminology below is used in, for example, *Multivariable Mathematics* by Williamson and Trotter.

standard	*alternative*
kernel	null-space
range	image
codomain	range

As before, we consider a function $L: \mathcal{D} \to \mathcal{W}$. (That is, $\mathcal{D}$ is the domain and $\mathcal{W}$ is the codomain.) We'll use the notation appropriate to the case that $\mathcal{D}$ and $\mathcal{W}$ are vector spaces and L is linear — although the definitions of "range", "onto", and "one-to-one" also apply in more general contexts.

Range

Definition: The set of all values which $L(\vec{x})$ *actually attains* as $\vec{x}$ varies through all of $\mathcal{D}$ is the *range of* L. It is abbreviated $\operatorname{ran} L$, and also is sometimes symbolized by $L(\mathcal{D})$ or $L[\mathcal{D}]$ (see the definition of "image" below).

Examples:

1. If $\mathcal{D} = \mathbf{R}^2$, $\mathcal{W} = \mathbf{R}^2$, and

$$L\begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} y \\ x \end{pmatrix},$$

   then $\operatorname{ran} L$ is all of $\mathcal{W}$, since every $\binom{y}{x}$ can be reached by choosing the right input, $\binom{x}{y}$.

2. If $\mathcal{D} = \mathbf{R}^2$, $\mathcal{W} = \mathbf{R}^2$, and

$$L\begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} y \\ -y \end{pmatrix},$$

   then $\operatorname{ran} L$ is the subspace of vectors of the form $\binom{t}{-t}$ (which is not all of $\mathcal{W}$).

Important remark: The range of L is the set of $\vec{y} \in \mathcal{W}$ such that the equation

$$L(\vec{x}) = \vec{y}$$

has a solution $\vec{x} \in \mathcal{D}$.

As appliers of mathematics, this is why we are interested in the range of a function! It gives a new way of thinking and talking about the very practical question of solving equations. We cannot overemphasize that the idea of range is not something new and strange, but rather a formalization or codification of something very familiar, the question of whether a given equation is solvable. The advantage of using the new abstract terminology is at least twofold: (1) It gives access to the theorems and discussions in the mathematical literature which are written in this language. (2) More fundamentally, it concentrates the mind on the *function* as a real object, a mapping from one set into another, as opposed to the equation as a string of marks on paper. This is a major step on the path to mathematical wisdom. This enlightened point of view is essential for solving problems that don't yield to the routine application of memorized rules for manipulation of symbols.

Example: Consider the differential equation

$$\frac{dy}{dt} = f(t) \quad \text{with the boundary conditions} \quad y(0) = 0 = y(1).$$

We can think of the operator $L = d/dt$ as acting on the domain $\mathcal{D}$ consisting of all differentiable functions on the interval $(0, 1)$ that vanish at both endpoints; L takes its values in the codomain $\mathcal{W} = \mathcal{C}(0, 1)$. The only solution satisfying the left-end boundary condition is

$$y(t) = \int_0^t f(\tilde{t})\, d\tilde{t}.$$

It will satisfy the right-end boundary condition if and only if

$$\int_0^1 f(t)\, dt = 0. \qquad (\#)$$

The range of L (with the restricted domain discussed here!) consists of just those special elements of $\mathcal{W}$ that satisfy (#).

Theorem 1: If L is linear, then its range is a subspace (not just a subset) of the codomain $\mathcal{W}$.

PROOF: If $\vec{y}_1 = L(\vec{x}_1)$ and $\vec{y}_2 = L(\vec{x}_2)$ for two elements $\vec{x}_1$ and $\vec{x}_2$ in $\mathcal{D}$, then for any $r \in \mathbf{R}$,

$$r\vec{y}_1 + \vec{y}_2 = rL(\vec{x}_1) + L(\vec{x}_2) = L(r\vec{x}_1 + \vec{x}_2).$$

Therefore, $r\vec{y}_1 + \vec{y}_2$ is in the range of L, so the range is algebraically closed (is a subspace).

Important remark: This theorem and its proof are a generalization of something you learn in your first course on differential equations. Namely, if you know how to solve an nonhomogeneous linear differential equation, $L[y] = g$, for two choices of the "right-hand side", g_1 and g_2 , then you can solve the equation when g equals any linear combination of g_1 and g_2 — you just add the known solutions together with the same numerical coefficients:

$$L[y_i] = g_i \;\Rightarrow\; L[c_1y_1 + c_2y_2] = c_1g_1 + c_2g_2\,.$$

Our theorem is the same thing, except for differences in notation ($\vec{x}$ instead of y, etc.) and for the fact that we discard the completely irrelevant assumption

that the linear operator is *differential*. At root, then, the "superposition principle" for solving nonhomogeneous differential equations is a fact about vector spaces, not about differential equations! Once again, theorems like this *tell you how to solve problems*, however remote from "the real world" they may appear at first glance.

Theorem 2: If $L\colon \mathbf{R}^n \to \mathbf{R}^p$ is linear with matrix A, then the range of L is the span of the columns of A.

PROOF: This is almost obvious from the definition of the matrix product $A\vec{x}$. (Look at how you would calculate a 2×2 numerical example.) If $\vec{y} = L(\vec{x}) = A\vec{x}$, then

$$y_i = \sum_{j=1}^{n} x_j A_{ij} = \sum_{j=1}^{n} x_j [L(\vec{e}_j)]_i \,.$$

That is,

$$\vec{y} = \sum_{j=1}^{n} x_j L(\vec{e}_j),$$

a linear combination of the columns of A.

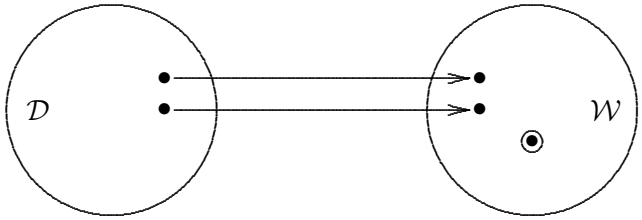
Definitions: If $\vec{x} \in \mathcal{D}$, then $L(\vec{x})$ $(\in \mathcal{W})$ is the *image of* $\vec{x}$ *under* L; and if $\mathcal{S}$ is a subset of $\mathcal{D}$, then the set of all vectors $L(\vec{x})$ for $\vec{x}$s in $\mathcal{S}$ is the *image of* $\mathcal{S}$ *under* L, denoted $L(\mathcal{S})$ or $L[\mathcal{S}]$. (For this reason, $\operatorname{ran} L$ is sometimes called the *image of* L and denoted by $L(\mathcal{D})$, as previously mentioned.)

For example, if $L\binom{x}{y} = \binom{y}{x}$ and $\mathcal{S}$ is the line $\binom{t}{0}$, then $L[\mathcal{S}]$ is the line $\binom{0}{t}$.

Theorem 1 generalizes to show that *the image under* L *of any subspace* $\mathcal{S}$ *of the domain* $\mathcal{D}$ *is a subspace of* $\mathcal{W}$.

ONE-TO-ONE AND ONTO

Definition: If the range of $L\colon \mathcal{D} \to \mathcal{W}$ is *all* of $\mathcal{W}$, then L is said to be a function from $\mathcal{D}$ *onto* $\mathcal{W}$. That is, the following type of situation is forbidden:



For short, when there is no doubt as to what space is playing the role of $\mathcal{W}$, one just says "L is onto". If you don't like turning prepositions into adjectives, an alternative term is *surjective*.

Important remark: L is onto if and only if the equation $L(\vec{x}) = \vec{y}$ can be solved for every $\vec{y} \in \mathcal{W}$ (the solution being in $\mathcal{D}$). (Here, "can be solved" means that a solution *exists*, not that anybody necessarily knows how to find it or write it down.)

There is a companion notion:

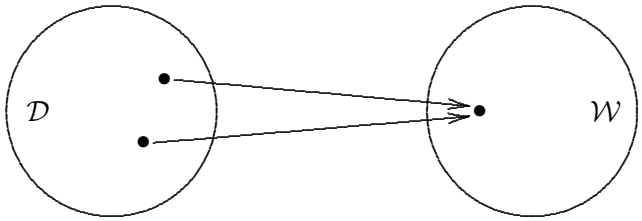
Definition: $L\colon \mathcal{D} \to \mathcal{W}$ is **not** *one-to-one* if there are two (or more) distinct points, $\vec{x}$ and $\vec{y}$, in $\mathcal{D}$ such that $L(\vec{x}) = L(\vec{y})$. That is, L **is** *one-to-one* (or *injective*) if

$$L(\vec{x}) = L(\vec{y}) \quad \text{implies} \quad \vec{x} = \vec{y},$$

or, equivalently,

$$\vec{x} \neq \vec{y} \quad \text{implies} \quad L(\vec{x}) \neq L(\vec{y}).$$

In other words, the following type of situation is forbidden:



Important remark: L is one-to-one if and only if the equation $L(\vec{x}) = \vec{y}$ never has more than one solution $\vec{x}$ in $\mathcal{D}$ for a given $\vec{y} \in \mathcal{W}$. That is, the solution is *unique*.

Kernel

Definition: If L is a function on $\mathcal{D}$ into $\mathcal{W}$, then the *kernel* of L is the set of $\vec{x}$s in $\mathcal{D}$ such that $L(\vec{x}) = \vec{0}$. (Note that this $\vec{0}$ is the zero vector of $\mathcal{W}$.)

Arguably even this definition makes sense for any function, not necessarily a linear one. For example, if L were a polynomial, the elements of its kernel (more often called *null set* in this context) would be just what we usually call the roots of the polynomial. For a linear function, however, the kernel has a special significance, as we'll see; furthermore, whenever the kernel of a linear function has more than one member, it has infinitely many, making up a subspace (rather than a discrete set of roots).

The kernel is denoted by $\ker L$ (also by $\mathcal{N}[L]$, especially when the alternative term "null-space" is used).

Examples:

1. Let $L\binom{x}{y} = \binom{y}{x}$. Then $\ker L = \{\vec{0}\} \equiv \left\{\binom{0}{0}\right\}$, since no vector other than $\vec{0}$ is mapped into $\vec{0}$ by L. In fact, this L is one-to-one; the output, $\binom{y}{x}$, uniquely determines the input, $\binom{x}{y}$.
2. Let $L\binom{x}{y} = \binom{y}{-y}$. Then its kernel consists of all vectors of the form $\binom{x}{0}$. Clearly, this function is not one-to-one.

The connection between injectivity and the kernel, observed in this example, is actually universal (for linear functions):

Theorem 3: If L is linear, then L is one-to-one if and only if its kernel is $\{\vec{0}\}$. (This $\vec{0}$ is the zero vector of $\mathcal{D}$. Notice that $\ker L$ always *contains* $\vec{0}$. The point of the theorem is that it contains *nothing else* if L is one-to-one.)

Proof: If $\ker L$ contains more than one vector, then L is not one-to-one, by definition. The converse is less immediate, but easy: If L is not one-to-one, we have $L(\vec{x}) = L(\vec{y})$ but $\vec{x} \neq \vec{y}$, for some $\vec{x}$ and $\vec{y}$. By linearity,

$$L(\vec{x} - \vec{y}) = L(\vec{x}) - L(\vec{y}) = 0\,[\in \mathcal{W}].$$

Thus $x - y$ is a nonzero element of $\ker L$.

The trick used in this proof is worth remembering, because it is also useful in applications. We convert a question about an *nonhomogeneous* equation, $L(\vec{x}) = \vec{y}$, into a question about the corresponding *homogeneous* equation, $L(\vec{x}) = 0$, which is usually easier to analyze. (In the proof, the "question" was about uniqueness of solutions.)

Theorem 4: If L is linear, then $\ker L$ is a subspace of $\mathcal{D}$.

PROOF: If $L(\vec{x}) = 0$ and $L(\vec{y}) = 0$, then $L(r\vec{x} + \vec{y}) = rL(\vec{x}) + L(\vec{y}) = 0$.

Important remark: $\ker L$ is the space of solutions (in $\mathcal{D}$) of the homogeneous equation, $L(\vec{x}) = 0$.

Examples of homogeneous equations:

1. An algebraic equation ($\mathcal{D} = \mathcal{W} = \mathbf{R}^2$): $\begin{cases} 2x + 3y = 0 \\ x - y = 0 \end{cases}$

   This system is nonsingular, so $\ker L = \{0\}$.

2. An ordinary differential equation: $\dfrac{d^2 f}{dt^2} + 4f = 0.$

   This equation has many solutions, so $\ker L \neq \{0\}$. (Note, incidentally, that the f in this example is playing the role of the $\vec{x}$ in our general discussion. The linear function (L in the general discussion) here is the linear operator $(d/dt)^2 + 4$, which maps functions (the vectors of this example) into functions.)

3. A partial differential equation: $\nabla^2 f \equiv \dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} = 0.$

   (Again, the vectors are the functions f; here x and y play the role of t in the previous example.) This also has many solutions ($\ker L \neq 0$). However, we could impose the additional *boundary condition* that $f(x, y) = 0$ for x and y on the boundary of a bounded region in $\mathbf{R}^2$ (and consider the functions f to be defined only inside that region). The boundary condition is also a homogeneous linear equation; it goes together with the PDE (Laplace's equation) to make a homogeneous linear "system" like the two algebraic equations in the first example. It can be shown that the only solution of this system or problem as a whole is $f \equiv 0$. Thus $\ker L = \{0\}$ for Laplace's equation with the boundary condition included.

Important remark: The solutions of an *nonhomogeneous* linear equation, $L(\vec{x}) = \vec{y}$, form an *affine subspace* — namely, $\ker L + \vec{p}$, where $\vec{p}$ is any "particular solution" (i.e., $L(\vec{p}) = \vec{y}$). (Of course, $\vec{y}$ must be in the range of L in order for any solutions to exist.)

The proof of the principle expressed in this remark is the same as that for the special case taught in ODE courses: Given a particular solution $\vec{p}$, we have $L(\vec{x} + \vec{p}) = L(\vec{x}) + L(\vec{p}) = L(\vec{x}) + \vec{y}$. Therefore, $L(\vec{x} + \vec{p}) = \vec{y}$ if and only if $L(\vec{x}) = 0$ — i.e., $\vec{x} \in \ker L$.

Examples of nonhomogeneous equations:

1. An algebraic equation: $\begin{cases} 2x + 3y = 2 \\ x - y = -3 \end{cases}$
2. An ordinary differential equation: $\dfrac{d^2f}{dt^2} + 4f = e^t$.
3. Some partial differential equations with boundary conditions:
   (a) $\nabla^2 f = \rho(x, y)$ (known, nonzero); $f(x, y) = 0$ on the boundary.
   (b) $\nabla^2 g = 0$; $g(x, y) = \gamma(x, y)$ (a known, nonzero function) on the boundary.

   In both cases the solution of the nonhomogeneous problem is unique because $\ker L = \{0\}$ for the corresponding homogeneous problem (which was the third in our previous list of examples). Notice that $h = f + g$ solves the problem with both nonhomogeneities,

$$\nabla^2 h = \rho, \qquad h = \gamma \quad \text{on the boundary.}$$

Summary examples

Example 1. Let $\mathcal{D} = \mathcal{W} = \mathcal{P}_3$, the vector space of polynomials of degree 3 or less:

$$p(t) = at^3 + bt^2 + ct + d.$$

Define $L\colon \mathcal{P}_3 \to \mathcal{P}_3$ by

$$L(p) \equiv p' \quad \text{(i.e., } L \equiv d/dt\text{)}.$$

Calculate:

$$[L(p)](t) = 3at^2 + 2bt + c. \tag{†}$$

The *kernel* of L is the set of ps that get mapped into the zero polynomial; these are just the constant polynomials, so

$$\ker L = \mathcal{P}_0 \equiv \{d\}.$$

Please don't confuse $\ker L$ with the set of roots of the polynomial $3at^2 + 2bt + c$ for fixed values of a, b, c! Remember that the vectors in this problem are polynomials, so the kernel must be a set of polynomials, not a set of numbers. The *range* of L is $\mathcal{P}_2$, since every possible quadratic polynomial

is obtained for some choice of the constants in (†). [In more detail: If $L(p)$ is $At^2 + Bt + C$, then we can solve for the coefficients of a corresponding p: $a = \frac{1}{3}A$, $b = \frac{1}{2}B$, $c = C$, d arbitrary.] This linear function is not *one-to-one*, since $\ker L \neq \{0\}$. Neither is it *onto*, since $\operatorname{ran} L \neq \mathcal{W}$. Note, however, that we could have studied the same function with $\mathcal{P}_2$ as codomain instead of $\mathcal{P}_3$; in that case, L would be onto.

Example 2. We return to Example 6 in Sec. 4.4. The linear operator $L\colon \mathcal{P}_2 \to \mathcal{P}_2$ is defined by

$$[L(p)](t) = (t^2 - 4)p''(t) + tp'(t) - 4p(t).$$

Is L surjective (onto $\mathcal{P}_2$)? If not,* what is its range? Is L injective? If not, what is its kernel?

SOLUTION: The function is not *surjective*; the *range* is $\mathcal{P}_1$, the first-degree polynomials. This is clear either from the matrix found in Sec. 4.4 — the span of the columns being the same as the span of

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

— or from the action of L on the basis polynomials, which yields the constants and the multiples of t but no t^2 terms. Neither is it *injective*: To see what the kernel is, reduce the matrix:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ -8 & 0 & -4 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus $p(t) = at^2 + bt + c$ is in the kernel if

$$a + \tfrac{1}{2}c = 0, \qquad b = 0.$$

That is, the kernel consists of the multiples of the polynomial

$$-\tfrac{1}{2}t^2 + 1.$$

We'll have still more to say about this example in Secs. 5.3 and 5.4.

---

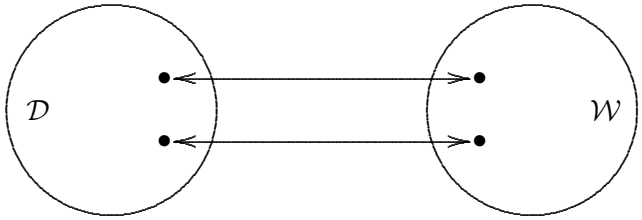
* If L is surjective, then its range is $\mathcal{P}_2$ and there is nothing left to ask about the range. Similarly, if L is injective, then its kernel is $\{\vec{0}\}$ and there is nothing more to say.

Isomorphisms

Definition: $L\colon \mathcal{D} \to \mathcal{W}$ is *bijective* (or is *a one-to-one correspondence between* $\mathcal{D}$ *and* $\mathcal{W}$) if L is both injective and surjective (i.e., both one-to-one and onto).

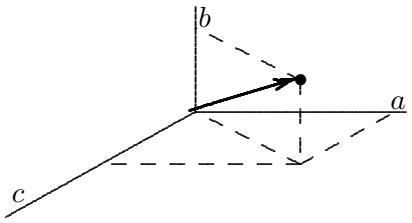
Definition: L is an *isomorphism* if it is both linear and bijective. (That is, L is linear, $\operatorname{ran} L$ equals all of $\mathcal{W}$, and $\ker L = \{0\}$.)

If $L\colon \mathcal{D} \to \mathcal{W}$ is an isomorphism, one says that $\mathcal{D}$ and $\mathcal{W}$ are *isomorphic*. The important point is that for many purposes we may then think of $\mathcal{D}$ and $\mathcal{W}$ as being *the same space*, since their internal algebraic structure is the same.



Example: $\mathcal{P}_2$, $\mathbf{R}^3$, and 3-dimensional physical space are all isomorphic:

$$at^2 + bt + c \;\leftrightarrow\; \begin{pmatrix} a \\ b \\ c \end{pmatrix} \;\leftrightarrow$$



Remark: A function is bijective if and only if it has an inverse function (see end of Sec. 3.2). Therefore, a linear function between finite-dimensional vector spaces is an isomorphism if and only if its matrix† is nonsingular (see Chapter 2). (This implies, in particular, that $\mathcal{D}$ and $\mathcal{W}$ have the same dimension. That is no surprise, since we already remarked that they are practically the same space in this situation. There will be more on this in Sec. 5.4, along with discussion of what can happen when the two dimensions are *not* equal.)

† The matrix of L depends on the bases chosen for $\mathcal{D}$ and $\mathcal{W}$. The statement, however, is true for any of these matrices.

Exercises

5.2.1 Consider the linear system

$$\begin{aligned} x + 2y + 2z + 2w &= 3, \\ 2x + 4y + 5z + 3w &= 8. \end{aligned}$$

(a) Find all solutions of the system.

(b) The coefficient matrix of the system, $M = \begin{pmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 5 & 3 \end{pmatrix}$, defines a linear function from $\mathbf{R}^4$ into $\mathbf{R}^2$. Describe the kernel and the range of this function.

(c) Explain how part (a) provides an example of the concept of "affine subspace".

5.2.2 A linear function F is defined in the usual way by the matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

(a) Find a basis for the range of F.

(b) Find a basis for the kernel of F.

(c) What are the dimensions of the subspaces associated with F — its domain, codomain, range, and kernel?

5.2.3 The equations $\left\{ \begin{array}{l} u = x - 2y \\ v = x + 3y \\ w = x - 6y \end{array} \right\}$ define a linear function $L\colon \mathbf{R}^2 \to \mathbf{R}^3$.

(a) Find the kernel and range of L.

(b) Is L injective (one-to-one)? Explain.

5.2.4 The matrix $A = \begin{pmatrix} 2 & 3 & 2 \\ -1 & 1 & 1 \end{pmatrix}$ defines a linear function $L\colon \mathbf{R}^3 \to \mathbf{R}^2$.

(a) Find the kernel of L.

(b) Is L "onto" (surjective)? Explain.

5.2.5 The function

$$F \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - 3x_2 \\ x_1 + x_2 \end{pmatrix}$$

is a linear mapping from $\mathbf{R}^2$ to $\mathbf{R}^2$. Find its kernel and its range.

5.2.6 Find the kernel and the range of the linear operator $L\colon \mathbf{R}^4 \to \mathbf{R}^4$ whose matrix is

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ -2 & -4 & 0 & -6 \\ 10 & 20 & 0 & 30 \end{pmatrix}.$$

5.2.7 Consider the linear function $L\colon \mathcal{P}_2 \to \mathcal{P}_2$ defined by $[L(p)](t) = p'(t) + tp''(t)$.

(a) What is the matrix representing L with respect to the traditional basis for $\mathcal{P}_2$, $\{t^2, t, 1\}$?

(b) Is L surjective? If not, what is its range, and what is the dimension of the range?

(c) Is L injective? If not, what is its kernel, and what is the dimension of the kernel?

5.2.8 Find the range and the kernel of each of these operators.

(a) $L\colon \mathcal{P}_2 \to \mathcal{P}_3\,, \quad L(p)(t) \equiv t\,p(t)$

(b) $L\colon \mathcal{P}_2 \to \mathcal{P}_2\,, \quad L(p)(t) \equiv t\,p''(t)$

5.2.9 Define $K\colon \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ by $K(f)(t) = t^2 f(t) + \int_0^t f(\tilde{t})\,d\tilde{t}$. This function is linear (Exercise 3.2.11). Is K surjective (onto)? HINT: What is $K(f)(0)$?

5.2.10 Define $L\colon \mathcal{P}_2 \to \mathcal{P}_2$ by $[L(p)](t) = p''(t) + (t+1)p'(t) - p(t)$.

(a) Find the matrix representing L with respect to the basis $\{t^2, t, 1\}$.

(b) Find the kernel of L.

(c) Find the range of L.

5.2.11 Define $K\colon \mathcal{P}_2 \to \mathcal{P}_2$ by $K(p) = p'' + (3+t)p' + p$.

(a) Find the matrix representing K with respect to the usual basis, $\{t^2, t, 1\}$.

(b) Find the range of K.

(c) Find the kernel of K.

5.2.12 Define a function $G\colon \mathcal{C}(-\infty,\infty) \to \mathcal{C}(-\infty,\infty)$ by

$$[Gf](x) \equiv \int_0^x t\, f(t)\, dt.$$

(a) Prove that G is linear.

(b) Show that G is injective.

(c) Show that G is *not* surjective. (In other words, find an element of $\mathcal{C}(-\infty,\infty)$ that is not equal to Gf for any f.)

5.2.13 Consider the function $L\colon \mathcal{P}_3 \to \mathcal{P}_3$ defined by $L(p) \equiv p' + p$.

(a) Show that L is linear (as a mapping of polynomials into polynomials).

(b) Is L surjective (onto)? If not, describe the range of L.

(c) Is L injective (one-to-one)? If not, describe the kernel of L.

5.2.14 Let $\mathcal{M}$ be the vector space of 2×2 matrices (cf. Exercise 5.1.4). Let A be a fixed element of $\mathcal{M}$ and X be a variable standing for elements of $\mathcal{M}$.

(a) Show that the formula $\quad G(X) \equiv AX \quad$ defines a linear function $G\colon \mathcal{M} \to \mathcal{M}$. (NOTE: This is not the same thing as the linear function from $\mathbf{R}^2$ to $\mathbf{R}^2$ represented by AX for a particular matrix X.)

(b) What is the range of G, if A is nonsingular? (Again, don't confuse this with the range of AX.)

(c) What is the range of G if $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$?

5.2.15 What can you say about the kernel of the function G defined in the preceding exercise? HINT: When does the range of [the linear function represented by] X match up with the kernel of [the linear function represented by] A?

5.2.16 Answer parts (b) and (c) of Exercise 5.2.14 for the case where the fixed matrix is on the right: $G(X) \equiv XA$.

In the remaining exercises, determine [bases for] the kernel and the range of the indicated exercise from Sec. 4.4.

5.2.17 Exercise 4.4.3.

5.2.18 Exercise 4.4.5.

5.2.19 Exercise 4.4.8.

5.2.20 Exercise 4.4.13.

5.3 Linear Equations: The Superposition Principles

The preceding section was decorated with paragraphs labeled "Important Remark", which indicated why each of the abstract definitions of that section is important for applications. Indeed, those remarks are so important that we shall now state them again in a different way, organized from the point of view of the applications. We have in mind primarily the solution of differential equations, ordinary and partial. The techniques presented in differential-equations courses sometimes strike students as simply a bewildering welter of magic tricks. However, many of them become well motivated, even obvious, once they are understood as general principles of linear algebra.

We begin by restating the obvious.

Definition: A *linear equation* is an equation of the form

$$L(\vec{u}) = \vec{g},$$

where L is a linear operator, $\vec{g}$ is a "given" or "known" vector, and $\vec{u}$ is the unknown vector to be solved for.

Here the vectors $\vec{g}$ and $\vec{u}$ may be vectors in the elementary sense (elements of $\mathbf{R}^n$ spaces), numbers (elements of $\mathbf{R}$), functions (elements of function spaces such as $\mathcal{C}^2(0,1)$), or elements of any other vector spaces. The two vectors do not need to be of the same type. For any linear operator $L\colon \mathcal{D} \to \mathcal{W}$, we can consider the equation $L(\vec{u}) = \vec{g}$ with $\vec{u}$ in the domain $\mathcal{D}$ and $\vec{g}$ in the codomain $\mathcal{W}$. Normally in an application the equation comes first, and a formal specification of the domain and codomain comes later in order to make the problem statement precise.

When a particular linear equation is written down, such as

$$u_1 + 2u_2 = 0,$$
$$u_1 - 3u_2 = 1,$$

or

$$\frac{du}{dt} + t^3\, u = \cos 3t,$$

or

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\,,$$

$\vec{g}$ or its components are recognizable as the terms that "don't depend on $\vec{u}$". (The other terms make up $L(\vec{u})$. Since the equation is linear, they involve components of $\vec{u}$ *exactly to the first power*.) It is customary to write the $\vec{u}$-dependent terms on the left side of the equation and the $\vec{u}$-independent terms on the right, but since any term could be added to or subtracted from both sides of the equation without changing its solutions, this convention is arbitrary; indeed, we violated it in the last of the three examples above. That example is a partial differential equation (PDE), the one before it was an ordinary differential equation (ODE), and the first example was an algebraic equation with domain and codomain $\mathbf{R}^2$.

Homogeneous vs. nonhomogeneous equations

Our example PDE, $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ (called the elementary *heat equation*), does not contain any $\vec{u}$-independent terms. This represents an absolutely crucial distinction:

Definition: A linear equation, $L(\vec{u}) = \vec{g}$, is *homogeneous* if $\vec{g} = \vec{0}$ (i.e., *all* terms in the equation are *exactly* of the first degree in $\vec{u}$); it is *nonhomogeneous* if $\vec{g} \neq \vec{0}$ (i.e., "constant" terms also appear).

In the second parenthetical clause, "constant" means independent of $\vec{u}$. The "constant" term $\vec{g}$ may be a nontrivial function of the *independent* variable(s) of the problem, as in our example ODE. That equation is nonhomogeneous (because of the $\cos 3t$), and the PDE example is homogeneous. The algebraic example is nonhomogeneous because of the 1. Here we are thinking of that system of simultaneous equations as a *single* linear equation in which the unknown quantity is a two-component vector,

$$\vec{u} \equiv \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

The linear operator L maps $\vec{u}$ onto another vector,

$$\vec{g} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In solving a differential equation one usually needs to deal with initial or boundary conditions in addition to the equation itself. The main reason is that initial or boundary data need to be specified to give the problem a unique answer. Usually these conditions are themselves linear equations — for example, a standard initial condition for the heat equation is

$$u(0, x) = f(x),$$

where f is some given function. Often the differential equation will be homogeneous but at least one of the boundary conditions will be nonhomogeneous. (The reverse situation also occurs.) We can think of the differential equation and the supplementary conditions as fitting together into a system of "simultaneous linear equations" just like the two halves of the 2×2 linear system. Therefore, let us introduce one more bit of jargon:

Definitions: A *linear problem* consists of one or more linear conditions (equations) to be satisfied by the unknown, $\vec{u}$. A linear problem is *homogeneous* if all of its conditions are homogeneous, *nonhomogeneous* if one or more of the conditions are nonhomogeneous.

Example A: The ODE problem

$$u'' + 4u = 0, \qquad u(0) = 1, \quad u'(0) = 0$$

is an nonhomogeneous linear problem (but based on a homogeneous ODE).

Example B: The PDE problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + j(x), \qquad u(0, x) = 0, \quad u(t, 0) = 0, \quad u(t, 1) = 0$$

is an nonhomogeneous linear problem. The boundary conditions *and* the initial condition are homogeneous, but the heat equation itself is nonhomogeneous in this case. (The function j represents generation of heat inside the conducting body, perhaps by combustion or radioactivity.)

Remark: It is easy to see that every homogeneous linear equation has $\vec{u} = \vec{0}$ as a solution. (One proof: $L(\vec{0}) = L(\vec{u} - \vec{u})$ (for any $\vec{u}$) $= L(\vec{u}) - L(\vec{u}) = \vec{0}$, QED.) Therefore, any homogeneous linear problem has $\vec{0}$ as a solution. Therefore, *if a linear problem has a unique solution and that solution is nontrivial (not just the zero vector), then that linear problem must be nonhomogeneous.* That is, an interesting, well-posed problem always has at least one nonhomogeneous condition.

Solving linear problems

The importance of linear problems is that solving them is made easy by the *superposition principles* (which don't apply to nonlinear problems):

Principles of Superposition:

1. *A linear combination of solutions of a homogeneous problem is a new solution of that problem.* That is, if $L(\vec{u}_1) = 0$ and $L(\vec{u}_2) = 0$, then $L(c_1\vec{u}_1 + c_2\vec{u}_2) = 0$ for any numbers c_1 and c_2 (and similarly for more than two solutions, and for more than one homogeneous linear equation defining the problem).

Example: Let Problem 1 be the homogeneous ordinary differential equation $u'' + 4u = 0$. Two solutions of this problem are

$$u_1 \equiv \cos 2t, \qquad u_2 \equiv \sin 2t.$$

Then $u = u_1 + 3u_2$, for example, is also a solution. (In fact, we know that the *most general* solution is $c_1u_1+c_2u_2$ where the cs are arbitrary constants. But for this we need a deeper existence-and-uniqueness theorem for second-order ODEs; it doesn't just follow from linearity.)

2. *The sum of a solution of an nonhomogeneous problem and a solution of the* corresponding homogeneous problem *is a new solution of the original nonhomogeneous problem.* "Corresponding homogeneous problem" means the one with the same L (i.e, the same "left-hand sides" in all equations) but with all $\vec{g}$ components replaced by 0.

Example: Let Problem 2 be the nonhomogeneous equation $u'' + 4u = e^t$. One solution is $u_{\mathrm{p}} \equiv \frac{1}{5}e^t$. (This has to be found by the method of undetermined coefficients, or by luck. Again, general principles of linearity by themselves can't solve the whole problem.) Now if we add a solution of Problem 1 we get a new solution of Problem 2: $u_3 \equiv \frac{1}{5}e^t + \cos 2t$.

3. *The difference of two solutions of an nonhomogeneous problem is a solution of the corresponding homogeneous problem.* Therefore (as a corollary of this principle and the preceding one together), *every* solution of an nonhomogeneous problem can be obtained from *one particular* solution of that problem by adding some solution of the homogeneous problem.

EXAMPLE: The general solution of Problem 2 is

$$u = \tfrac{1}{5}e^t + c_1 \cos 2t + c_2 \sin 2t.$$

4. *The sum of solutions to two nonhomogeneous problems with the same L is a solution of a* NEW *nonhomogeneous problem, for which the $\vec{g}$ is the* SUM *of the $\vec{g}$s of the two original problems.* (This principle extends immediately to a list of more than two nonhomogeneous problems.)

EXAMPLE A: The sum of two solutions of Problem 2, $u_{\rm p}$ and u_3, is $z \equiv \frac{2}{5}e^t + \cos 2t$, which is a solution of $z'' + 4z = 2e^t$. The important lesson to be learned from this example is that the right-hand side of this new equation is *not* e^t, the nonhomogeneous term of the two old equations. *Do not superpose solutions of an* NONHOMOGENEOUS *problem in the hope of getting a solution of that* SAME *problem.*

EXAMPLE B: Note that $u_{\rm p}$ is the unique solution of PROBLEM 3:

$$u'' + 4u = e^t, \qquad u(0) = \tfrac{1}{5}, \quad u'(0) = \tfrac{1}{5}\,.$$

Suppose that we really want to solve PROBLEM 4:

$$u'' + 4u = e^t, \qquad u(0) = 0, \quad u'(0) = 0.$$

Recalling Principles 2 and 3 as applied to the differential equation alone (not the initial conditions), we see that $u = u_{\rm p} + y$, where y is some solution of $y'' + 4y = 0$. A moment's further thought shows that the correct y is the solution of PROBLEM 5:

$$y'' + 4y = 0, \qquad y(0) = -\tfrac{1}{5}, \quad y'(0) = -\tfrac{1}{5}.$$

A standard calculation shows that $y = -\frac{1}{5}\cos 2t - \frac{1}{10}\sin 2t$, and from this and $u_{\rm p}$ we can get the solution of Problem 4. (Of course, in solving such problems one usually doesn't write out Problem 5 as an intermediate step; the standard procedure is to impose the initial data of Problem 4 on the general solution found earlier. That is just a different way of organizing the same algebra. However, consciously splitting an nonhomogeneous problem into two nonhomogeneous problems, as demonstrated here for an ODE, is a standard and very important technique for solving PDEs.)

In summary, these principles provide the basic strategies for solving linear problems. If the problem is nonhomogeneous and complicated, you should split it into simpler nonhomogeneous problems and add the solutions. If the solution is not unique, the nonuniqueness resides precisely in the possibility of adding a solution of the corresponding homogeneous problem. (In particular, if the original problem is homogeneous, then you seek the general solution as a linear combination of some list of basic solutions.) If the problem statement contains enough initial and boundary conditions, the solution will be unique; in that case, the only solution of the homogeneous problem is the zero vector (of the domain).

An example of this strategy is the solution of the heat-conduction problem in a bar with fixed end temperatures. The physical process is governed by the partial differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

satisfied by the temperature function $u(t, x)$. To predict the temperature distribution in the future, we must know what it is now; so there is an *initial condition*,

$$u(0, x) = f(x).$$

Finally, to complete the physical description of the problem, one must specify what happens to the heat when it reaches the ends of the bar; here we are considering the case of *boundary conditions*

$$u(t, 0) = T_1\,, \qquad u(t, 1) = T_2\,.$$

(The end temperatures are *constants*, not functions of t as they could be in principle.) Here we have a homogeneous PDE, an nonhomogeneous initial condition (IC), and two nonhomogeneous boundary conditions (BC).

The details of how to solve this problem belong in another course. However, we can describe the basic trick. It is to treat the two types of nonhomogeneity separately. One writes $u = v + w$, where:

1. v is to be a solution of the problem that consists of the PDE and the nonhomogeneous BC, with *no* particular IC assumed. It is possible to find a solution of this problem that is *independent of* t (satisfies $\frac{d^2v}{dx^2} = 0$ and the BC):

$$v(t, x) = V(x) \equiv (T_2 - T_1)x + T_1\,.$$

2. w is to be a solution of the problem that consists of the PDE, the *homogeneous* boundary conditions

$$w(t,0) = 0, \qquad w(t,1) = 0,$$

and the initial condition that is needed to make u satisfy the original IC — namely,

$$w(0,x) = f(x) - V(x) \equiv g(x).$$

It is very important that the only nonhomogeneity in the problem is the IC. This makes it possible to solve for w by a method called *separation of variables* or *Fourier analysis*. In brief summary: First one observes that all the functions

$$w_n(t,x) \equiv e^{-n^2\pi^2 t}\sin(n\pi x) \qquad (n \text{ a positive integer})$$

satisfy both the PDE and the homogeneous BC. Then it turns out that the arbitrary given function g, if it satisfies some reasonable technical conditions, can be written as an infinite linear combination (*Fourier series*) of the functions $\sin(n\pi x)$. Because the PDE and BC are both homogeneous, it is possible to add up the solutions w_n without falling into the trap warned against in Example A under Principle 4 above; the result satisfies the PDE, BC, and IC required of w. (Thus $u \equiv v + w$ satisfies the original problem, and we're done!) The important point is that for this method to work, it is absolutely crucial to make the boundary conditions homogeneous first.

Summary of the relations among subspaces, linear functions, and linear equations

We have three different sets of terminology, depending upon whether the focus is on the subspace, the linear function, or the equation. Their correspondence is exhibited by the following table.

Property of key subspace	**Property of linear function**	**Property of linear equations**
kernel $= \{0\}$ (not bigger)	one-to-one (injective)	uniqueness (homog. & nonhomog.)
range $=$ codomain (not smaller)	onto (surjective)	existence (nonhomogeneous)

Nonhomogeneous equations are associated with affine subspaces and functions:

Type of linear equation	**Solutions are zero set of**	**Type of subset**
homogeneous, $L(\vec{x}) = 0$	linear function, $L(\vec{x})$	("true") subspace, $\ker L$
nonhomogeneous, $L(\vec{x}) = \vec{b}$	affine function, $L(\vec{x}) - \vec{b}$	affine subspace, $\ker L + \vec{p}$

More examples

Example 1. Use superposition principles to find all solutions in $\mathcal{P}_2$ of the differential equation

$$(t^2 - 4)p''(t) + tp'(t) - 4p(t) = -16.$$

Solution: We have already analyzed the linear operator

$$L = (t^2 - 4)\,\frac{d^2}{dt^2} + t\,\frac{d}{dt} - 4$$

with domain and codomain $\mathcal{P}_2$ in Example 6 of Sec. 4.4 and Summary Example 2 of Sec. 5.2. We saw that $L(t^2) = -8$; consequently, $L(2t^2) = -16$, so one solution of the equation is $p_{\mathrm{p}}(t) = 2t^2$. The general solution, therefore, is $p = p_{\mathrm{p}} + p_{\mathrm{c}}$, where p_{c} is the general element of the kernel (general solution of the corresponding homogeneous equation, $L(p) = 0$). We already identified the kernel in the "summary example". So we see that all the solutions of our equation in $\mathcal{P}_2$ are

$$p(t) = 2t^2 + \left(-\tfrac{1}{2}t^2 + 1\right) c \quad \text{for arbitrary numbers } c.$$

(There are other solutions of this differential equation that are not quadratic polynomials.)

Example 2. Recall from third-semester calculus or physics (or look ahead to Sec. 6.4) that ∇^2 stands for the *Laplacian operator*,

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\,.$$

Let S be the sphere $x^2+y^2+z^2=1$ and B be the ball $x^2+y^2+z^2<1$. We investigate the partial differential equation $\nabla^2 u = f$ when the (real-valued) functions u and f have B as their domain ($f(\vec{r})$ given but arbitrary, $u(\vec{r})$ unknown and to be found).

(1) What are a reasonable domain* and codomain for ∇^2? Because differentiation "roughens" functions, we can't take these two spaces to be the same. It is reasonable to expect the equation to have a solution for any f in $\mathcal{C}^0(B)$ (in English: any source term that is a continuous function inside the ball). On the other hand, $\nabla^2 u$ will not be defined unless the second derivatives of u exist. So we expect the solution u to be in $\mathcal{C}^2(B)$ (the space of twice-differentiable functions with domain B). In summary, we are studying the operator

$$\nabla^2 : \mathcal{C}^2(B) \to \mathcal{C}^0(B).$$

(2) Next, we'll observe that for a fixed f, the solutions of the equation $\nabla^2 u = f$ form an affine subspace of $\mathcal{C}^2(B)$, which is a "true" subspace if and only if f is the zero function. Unfortunately, abstract linear algebra by itself can carry one only so far; we need to assume without proof some true but harder theorems from the theory of differential equations. Namely: (1) Every solution of $\nabla^2 u = f$ is indeed a member of $\mathcal{C}^2(B)$ (if $f \in \mathcal{C}^0(B)$). (2) At least one such solution exists. Denote such a solution by u_{p}. If u is any other solution, then $v \equiv u - u_{\mathrm{p}}$ satisfies $\nabla^2 v = 0$, because ∇^2 is linear. Conversely, if v solves $\nabla^2 v = 0$ (called *Laplace's equation*), then $u = v + u_{\mathrm{p}}$ solves the original problem. (These arguments are just Superposition Principles 2 and 3 applied to this particular example.) In other words, the solution space is the affine subpace

$$\ker \nabla^2 + u_{\mathrm{p}}\,.$$

(If $f = 0$, then u_{p} itself lies in $\ker \nabla^2$ and we might as well choose it to be the zero function.)

(3) Again stepping outside the bounds of pure linear algebra, we note that there is a well known method for constructing a solution u_{p}. This is most familiar in the physics terminology†: The *Coulomb potential* associated

---

* Note that the domain of u and f is a subset of $\mathbf{R}^3$, and their codomain is $\mathbf{R}$, but the domain and codomain of ∇^2 are sets of *functions*. One needs to apply the abstract concepts at different levels in the same problem.

† But we shall ignore some subtleties about sign conventions and rival systems of units that need to be considered in the physics and engineering books.

with a *point charge* at $\vec{r}_0$ is $\frac{-e}{4\pi r}$ where e is the magnitude of the charge and $r \equiv \|\vec{r} - \vec{r}_0\|$ is the distance from the charge to the (variable) point $\vec{r}$. If we have a spread-out distribution of charge with density f, then the associated Coulomb potential is

$$u(\vec{r}) = \iiint \frac{-f(\vec{r}_0)}{4\pi\|\vec{r} - \vec{r}_0\|}\, d^3 r_0\,.$$

This means that $\nabla^2 u = f$. (This equation holds everywhere in $\mathbf{R}^3$, and therefore certainly inside B.) In the context of our problem, this Coulomb solution is a suitable candidate for $u_{\rm p}$.

(4) Suppose that we knew how to solve the problem

$$\nabla^2 u(\vec{r}) = 0 \quad \text{for } \vec{r} \in B, \qquad u(\vec{r}) = g(\vec{r}) \quad \text{for } \vec{r} \in S, \tag{\#}$$

for any given function $g(\vec{r}) = g(\theta, \phi) \in \mathcal{C}^0(S)$. We outline a strategy for solving the problem

$$\nabla^2 u(\vec{r}) = f(\vec{r}) \quad \text{for } \vec{r} \in B, \qquad u(\vec{r}) = g(\vec{r}) \quad \text{for } \vec{r} \in S.$$

This is an instance of Superposition Principle 4, with one nonhomogeneity living on the boundary sphere S and the other one inside B. From the previous discussion, we know that $u = u_{\rm p} + v$, where $u_{\rm p}$ is the Coulomb solution from Part 3 and v is some solution of the homogeneous Laplace equation. The slightly tricky part is the boundary condition satisfied by v. For $\vec{r} \in S$, we have

$$v(\vec{r}) = u(\vec{r}) - u_{\rm p}(\vec{r}) = g(\vec{r}) - u_{\rm p}(\vec{r}) \equiv h(\vec{r}).$$

Recall that the first term in h was prescribed in the statement of our problem. The second term is also known (provided that we can evaluate the integral defining $u_{\rm p}$). Therefore, to complete the solution for u we merely need to solve problem (#) with v in the role of u and h in the role of g.

Exercises

5.3.1 (a) Find the kernel of $L\colon \mathcal{C}^2(-\infty, \infty) \to \mathcal{C}(-\infty, \infty)$ defined by $L(u) = u'' - 4u$.

(b) What can you say about the range of L? Show that the range contains (at least) all bounded[‡] continuous functions. HINT: Construct the solution by the *method of variation of parameters*. The point of this problem is that almost every function is in the range, if we put no conditions on the behavior of the solution at infinity. We assume boundedness to guarantee that the integrals converge.

5.3.2 Define a linear operator $L: \mathcal{P}_1 \to \mathcal{P}_2$ by $[L(p)](t) \equiv p'(t) + t\,p(t)$.

(a) Find the kernel of L, as a finite-dimensional linear-algebra problem. (Use the power bases for the polynomial spaces.)

(b) Find the kernel of L, as a differential-equations problem. (Of course, your answer should agree with (a).)

(c) How would your answer to (b) change if we were considering

$$L: \mathcal{C}^2(-\infty, \infty) \to \mathcal{C}(-\infty, \infty),$$

defined by the same differential expression?

(d) Find the range of L.

5.3.3 Redo Example 1 using the observation that $L(1) = -4$ (instead of $L(t^2) = -8$). Comment on the consistency of your answer with the one in the text.

5.3.4 Consider *Poisson's equation*, $\nabla^2\phi = \rho$. (Physically, $\rho(\vec{r})$ is a given charge distribution and $\phi(\vec{r})$ is the electrical potential to be found.)

(a) Note that if ϕ is a solution, then $\phi + e^y \sin x$ is also a solution (with the same ρ). Explain what this has to do with the terms "injective" and "kernel".

(b) Suppose that the equation holds inside a sphere with center at the origin, and that the radial derivative $\frac{\partial\phi}{\partial r} = \hat{n} \cdot \nabla\phi$ (the directional derivative perpendicular to the sphere) equals 0 everywhere on the sphere. (This means on the surface, not the whole interior!) In Exercise 7.5.6 we shall show that the integral of ρ over the interior of the sphere must then equal 0. (Physically, the net charge is zero.) Explain what this phenomenon has to

‡ There is an M such that $|u(x)| < M$ for all x.

do with the terms "range" and "onto". HINT: Think of ∇^2 as a linear operator on a domain consisting of functions whose normal derivatives vanish on the sphere; see also the next exercise.

5.3.5 Here is a one-dimensional analogue of the phenomenon described in Exercise 5.3.4(b). Consider a steady-state heat-conduction problem similar to the one discussed in the text of this section, except that this time there is a heat source in the interior of the bar, and the derivatives of the temperature function at the endpoints are required to be 0. (Physically this means that there is no *heat flow* into or out of the bar at the ends; the ends are thermally insulated.) The differential equation is then

$$\frac{d^2u}{dx^2} = -j(x),$$

where j describes the heat source, and the boundary conditions are

$$\frac{du}{dx}(0) = 0, \qquad \frac{du}{dx}(L) = 0.$$

(a) Show that the problem has no solution unless $\int_0^L j(x)\,dx = 0$.

(b) Explain why this result should be expected on physical grounds. (Recall that u is a temperature function that is independent of time.)

(c) Let L be the operator d^2/dx^2 with the domain $\mathcal{D}$ consisting of those functions in $\mathcal{C}^2(0, L)$ that satisfy the thermal-insulation boundary conditions. Restate the result (a) as a description of $\operatorname{ran} L$.

(d) What is $\ker L$?

5.3.6 Find the general solution of

$$\frac{d^4y}{dx^4} - 16y = 0.$$

HINT: The fourth roots of 16 are ± 2 and $\pm 2i$.

5.3.7 Find the general solution of

$$\frac{d^4y}{dx^4} = \sin(3x).$$

5.4 **Rank**

For the past three chapters we have been suffering through a traffic jam of new concepts. Linear functions, linearly independent sets, and subspaces all need to be defined and explained one at a time, but it is hard to appreciate any of them fully without already understanding the others. This section finally pulls together all the pieces. We hope that it will make them all clearer.

Theorem 1: For any linear function,

	the dimension of the range
plus	the dimension of the kernel
equals	the dimension of the domain.

PROOF (for the case that all the spaces are finite-dimensional): Pick a basis $\{\vec{v}_1, \ldots, \vec{v}_p\}$ for the kernel. Add vectors to that basis to get a basis $\{\vec{v}_1, \ldots, \vec{v}_p, \vec{u}_1, \ldots, \vec{u}_q\}$ for the entire domain (Lemma, Sec. 5.1). The image of this set, $\{L(\vec{v}_1), \ldots, L(\vec{u}_q)\}$, must span the range. The images $L(\vec{v}_j)$ of the kernel vectors are all 0. The images $\{L(\vec{u}_j)\}$ of the extra vectors must be linearly independent: otherwise, some linear combination of the $\vec{u}$s would be in the kernel. (If $\sum_{j=1}^{q} c_j L(\vec{u}_j) = 0$, then $L\big(\sum_{j=1}^{q} c_j \vec{u}_j\big) = 0$.) So $\{L(\vec{u}_j)\}_{j=1}^{q}$ is a basis for $\operatorname{ran} L$. Thus we have

$$q = \dim \operatorname{ran} L, \quad p = \dim \ker L, \quad p + q = \dim \operatorname{dom} L,$$

which is the relationship to be proved.

Example: Look again at Summary Example 2 of Sec. 5.2,

$$L = (t^2 - 4)\,\frac{d^2}{dt^2} + t\,\frac{d}{dt} - 4$$

with domain and codomain $\mathcal{P}_2$. Once we found that the range ($\mathcal{P}_1$) has dimension 2, we could have concluded immediately that the kernel must have dimension $3 - 2 = 1$. That is the same conclusion we reached by brute-force calculation (which, of course, gave us the additional information of

exactly what one-dimensional subspace the kernel *is*). In particular, since the operator L was defined as a mapping from $\mathcal{P}_2$ into itself, once we saw that L is not surjective, we could have concluded that it is not injective either:

Corollary: If $\dim \mathcal{D} = \dim \mathcal{W} < \infty$, then $L: \mathcal{D} \to \mathcal{W}$ is one-to-one *if and only if* it is onto. (Either virtue guarantees the other.)

Notice that this theorem and corollary have a strong family resemblance to the earlier theorem relating independence to spanning for sets of various sizes in a space of a given dimension. The more a set fails to be independent (by being "flattened out"), the smaller is the subspace that it spans. Similarly, the larger the kernel of a linear function is, the smaller is its range.

In fact, Theorem 2 of Sec. 5.2 (together with the theory in Sec. 4.4) shows that ranges and spans are essentially the same thing (see the second definition in the next list).

Definitions:

(1) The *rank* of a *linear function* is the dimension of its range.

(2) The *rank* of a *matrix* is the rank of the linear function ($\mathbf{R}^n \to \mathbf{R}^p$) it represents. Equivalently, the rank is the dimension of the subspace of $\mathbf{R}^p$ spanned by the columns of the matrix. This number is often called "the number of linearly independent columns."

(3) The *nullity* of a linear function (or of the corresponding matrix) is the dimension of its kernel.

Example. What are the rank and the nullity of

$$M = \begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & 4 \end{pmatrix}?$$

SOLUTION: The *rank* is 2: It can't be greater than 2, because the dimension of the whole column space is just 2. It can't be less than 2, because the three columns are not all proportional. (Alternative argument: The (column) rank is equal to the row rank, and the row rank of this matrix is 2. A row reduction that shows this has already been done in Example 2(a) of Sec. 4.1.) Therefore, by Theorem 1, the *nullity* is $3 - 2 = 1$. This can also be checked directly, by reducing the augmented matrix for the homogeneous equation system,

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & -1 & 4 & 0 \end{pmatrix},$$

to

$$\begin{pmatrix} 1 & 0 & \frac{9}{5} & 0 \\ 0 & 1 & -\frac{2}{5} & 0 \end{pmatrix}.$$

From this reduced augmented matrix it is clear that the third coordinate of the solution is arbitrary and the first and second coordinates are then determined.

Application of the theorem to algebraic linear equations also known as Interpretation of the theorem in terms of algebraic linear equations

Theorem 1, relating nullity to rank, has a lot to teach us about the number of possible solutions of the homogeneous and nonhomogeneous systems associated with a matrix. You have probably known for a long time that a system *usually* has

- *many solutions* if there are more unknowns than equations,
- *no solutions* if there are more equations than unknowns,
- *a unique solution* if there are the same number of equations and unknowns.

On the other hand, you know that *there are exceptions* to these rules of thumb. With the concept of rank, we can get a much better handle on the types of behavior that can occur.

Let A be a $p \times n$ matrix, so that

$$A\vec{x} = \vec{y}$$

with $\vec{y}$ given is a system of p equations in n unknowns, representing a linear function from $\mathbf{R}^n$ to $\mathbf{R}^p$. Let M be the augmented matrix of the system, which is of shape $p \times (n+1)$; let M_{red} be its row-reduced form. The theorem states that the rank of A plus the nullity of A equals n:

$$r + q = n.$$

I. CASE $p = n$: (For illustrative purposes we take $p = 3$.) There are two possibilities:

A. A is invertible; $M_{\text{red}} = \left(\begin{array}{ccc|c} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & a_3 \end{array}\right)$. The equation has the *unique* solution $x_1 = a_1$, $x_2 = a_2$, ... This is the case of the theorem and corollary in which $q = 0$ (i.e., the solutions are unique) and $r = n = p$ (hence the solution exists for all $\vec{y}$; the range is the whole codomain).

B. A is singular. Then $q > 0$ (solutions are nonunique) and $r = n - q < p$ (for some $\vec{y}$s, solutions don't exist — i.e., $\vec{y}$ is not in the range). If $\vec{y}$ *is* in the range, then $A\vec{x} = \vec{y}$ has a *q-parameter family* of solutions,

$$\vec{x} = \vec{x}_{\text{particular}} + \text{arbitrary element of kernel.}$$

For $p = n = 3$, we can easily write down all possible kinds of results for M_{red}. For rank-2 matrices A, the most common result is

$$\left(\begin{array}{ccc|c} 1 & 0 & \alpha & a_1 \\ 0 & 1 & \beta & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right),$$

where the letters stand for arbitrary numbers. Solutions exist if and only if $a_3 = 0$ — which in turn is determined by the original coordinates of the nonhomogeneous term, $\{y_j\}$. Indeed, a_3 is some nontrivial linear combination of the three y_js. Therefore, two coordinates of $\vec{y}$ can be chosen arbitrarily, but then the third is determined, if $\vec{y}$ is to be in the range. This confirms that the dimension of the range (the rank) is 2. On the other hand, if a_3 *is* 0, then x_3 is arbitrary but the other x_js are determined. Thus the solution spaces are one-dimensional; the nullity is 1, in agreement with the theorem.

The other possible results for M_{red} when $r = 2$ are

$$\left(\begin{array}{ccc|c} 1 & \alpha & 0 & a_1 \\ 0 & 0 & 1 & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right) \quad \text{and} \quad \left(\begin{array}{ccc|c} 0 & 1 & 0 & a_1 \\ 0 & 0 & 1 & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right).$$

The foregoing remarks apply to them as well.

If $r = 1$, there are three possible reduced forms:

$$\left(\begin{array}{ccc|c} 1 & \alpha & \beta & a_1 \\ 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right), \qquad \left(\begin{array}{ccc|c} 0 & 1 & \alpha & a_1 \\ 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right),$$

$$\left(\begin{array}{ccc|c} 0 & 0 & 1 & a_1 \\ 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right).$$

Each a_j is some linear function of $\vec{y}$. In all these cases the range consists of all $\vec{y}$ for which $a_2 = a_3 = 0$; thus only one component of $\vec{y}$ is free to vary, so the range is one-dimensional. The theorem says that the nullity must be 2; this is so, because with only one nontrivial equation, two components of the solution may be chosen arbitrarily.

There is one other possible reduced form:

$$\left(\begin{array}{ccc|c} 0 & 0 & 0 & a_1 \\ 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & a_3 \end{array}\right).$$

This can happen only if A is the 0 matrix.

II. Now consider the case $p < n$ (fewer equations than unknowns). The rank (being the dimension of the range) is necessarily less than or equal to p (the dimension of the codomain). Therefore, $r < n$; and hence $q = n - r > 0$. (In fact, $q \geq n - p$.) Typically, therefore, the system will have nonunique solutions. (In particular, the homogeneous equation, $A\vec{x} = 0$, always has nontrivial solutions. By definition, q is the dimension of the solution space of that equation.)

A. A typical M_{red} is

$$\left(\begin{array}{ccc|c} 1 & 0 & \alpha & a_1 \\ 0 & 1 & \beta & a_2 \end{array}\right),$$

from which we see that at least one coodinate of the solution can be chosen arbitrarily. Here $r = p = 2$ and $q = n - p = 1$.

B. It is possible, however, to have $r < p$; the reduced matrix in such a case looks something like this:

$$\left(\begin{array}{ccc|c} 1 & \alpha & \beta & a_1 \\ 0 & 0 & 0 & a_2 \end{array}\right).$$

Here $r = 1$ and $q = 2 > n - p$. Thus for some $\vec{y}$s there are no solutions, while for the others there are even more solutions than in case A. (In no case is the solution unique when there are fewer equations than unknowns.)

III. Finally, suppose that $p > n$ (there are too many equations). Then $r = n - q \leq n < p$, so the range can't ever be the entire codomain. That is, for some $\vec{y}$ there are no solutions. A typical reduced augmented matrix looks like

$$\left(\begin{array}{cc|c} 1 & 0 & a_1 \\ 0 & 1 & a_2 \\ 0 & 0 & a_3 \\ 0 & 0 & a_4 \end{array}\right).$$

If the a_js next to the zero rows at the bottom are themselves 0, then solutions do exist. If the number of nonzero rows is n, the solution is unique; if it is less than n, the solution is not unique.

More about rank

If you have very sharp eyes, you may have noticed that in all the cases we analyzed in the foregoing discussion, the rank of the matrix was equal to the number of nonzero rows in A_{red} , although we defined it as the number of linearly independent columns of A. The equivalence of these two numbers follows from the next two theorems.

Theorem 2: The rank of A^{t} equals the rank of A. In other words, the dimension of the space spanned by the rows of A equals the dimension of the space spanned by the columns of A.

A direct proof of this theorem is rather messy. Nicer proofs use the concepts of inner product and orthogonality, although the theorem itself has nothing to do with them. Accordingly, we delay our proof to Chapter 8.*

Theorem 3: Elementary row operations do not change:

(A) the rank of the matrix;

(B) the space spanned by the rows of the matrix;

(C) the space spanned by the columns of the transpose of the matrix.

---

* Another such proof is given by G. Mackiw, A note on the equality of the column and row rank of a matrix, *Mathematics Magazine* **68** (1995), 285–286.

OUTLINE OF PROOF: Obviously (B) implies (A) and (C). (B) is true because row operations replace the rows by linear combinations of themselves, in a reversible way.

NOTE: Row operations *do* change the space spanned by the *columns* of the matrix (but not the dimension of that space, which is the rank).

Corollary: The rank of a matrix in row echelon form (that is, one that has been row-reduced) is equal to the number of nonzero rows in it.

OUTLINE OF PROOF: It's easy to show that the nonzero rows are linearly independent.

This theorem and corollary justify the row-reduction method for testing linear independence or constructing an independent set with the same span as a given set (method "(B)" in Sec. 4.1). You should also now be able to justify method "(A)".

Theorem 4: Let $L: \mathcal{V} \to \mathcal{V}$ be a linear function from a vector space into itself, and let L be represented (with respect to some basis for $\mathcal{V}$) by the (square) matrix A. Then the following conditions are equivalent:

(1) L (or A) has the *maximal rank* possible (namely, the dimension of $\mathcal{V}$).

(2) A is invertible (equivalently, bijective).

(3) The determinant of A is not zero.

Exercises

5.4.1 Find the rank of each of these matrices.

(a) $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 5 & 5 & 17 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 2 \\ 1 & 6 \end{pmatrix}$

5.4.2 Find the rank of each of these matrices.

(a) $\begin{pmatrix} 2 & -1 & 3 \\ 5 & 5 & -5 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & 2 \\ 3 & 1 & 2 \end{pmatrix}$

5.4.3

(a) Find all solutions of the system $\begin{cases} x + 2y - 3z = 2, \\ x - 2y + 4z = 1. \end{cases}$

(b) Find a basis for the range of the linear function whose matrix is $\begin{pmatrix} 1 & 2 & -3 \\ 1 & -2 & 4 \end{pmatrix}$.

(c) Comment on the relation between the number of parameters in the solution to (a) and the number of vectors in the solution to (b).

5.4.4 Let $M = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{pmatrix}$.

(a) What is the rank of M?

(b) What other ranks are possible for 2×3 matrices? Give an example of each.

(c) What is the dimension of the kernel of M?

5.4.5 The matrix $B = \begin{pmatrix} 2 & 3 \\ -1 & 1 \\ 1 & 1 \end{pmatrix}$ defines a linear function $L: \mathbf{R}^2 \to \mathbf{R}^3$.

(a) What is the rank of this matrix?

(b) What other ranks are possible for matrices of this shape (3×2)?

5.4.6 Define $f: \mathbf{R}^3 \to \mathbf{R}^2$ by $\begin{cases} u = \quad x - 3y + 2z, \\ v = -2x + 6y - 4z. \end{cases}$

(a) Is f one-to-one?

(b) Is f onto?

(c) Which of these two questions could have its answer changed by changing the numbers in the formulas, but leaving the spaces $\mathbf{R}^3$ and $\mathbf{R}^2$ the same? Explain.

5.4.7 Let M be a $p \times n$ matrix. The *index* of M is defined as the dimension of the kernel of M minus the dimension of the kernel of M^{t}. Prove that the index depends only on p and n, not on the numbers that make up M.

5.4.8 Find the rank and nullity of the operator in Exercise 4.5.8.

5.4.9 Find the rank and nullity of the operator in Exercise 4.5.14.

5.4.10 Prove that conditions (1) and (2) of Theorem 4 are indeed equivalent. (We postpone condition (3) to Chapter 7.)

5.4.11 Justify method "(A)" of Sec. 4.1 for testing linear independence. (You may appeal to Theorem 4, including the third condition.)

5.4.12 Prove: If L and M are linear transformations, then the rank of their product LM is less than or equal to the rank of either factor.

5.4.13 Unfinished business from Sec. 2.3: Prove that if A and B are square matrices satisfying $AB = I$, then A and B are invertible (and therefore $BA = I$ also). HINT: Use the result of the previous exercise.

5.5 Implicit and Inverse Functions

In one variable, the inverse function theorem is summarized in Leibnitz notation by

$$\frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1}.$$

It is no big surprise that the generalization to several variables will turn out to be

$$\frac{d\vec{x}}{d\vec{y}} = \left(\frac{d\vec{y}}{d\vec{x}}\right)^{-1} \qquad \text{(a matrix inverse)}.$$

Consider a function $\vec{y} = f(\vec{x})$, $f\colon \mathbf{R}^n \to \mathbf{R}^n$ (the *same* n), and suppose that a local inverse function (a "solution"), $\vec{x} = f^{-1}(\vec{y})$, is defined (at least for $\vec{y}$ in some region).

For example: Recall the *cylindrical coordinate* transformation,

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta, \\ z &= z. \end{aligned}$$

Here the relation between the general notation and that of the application is

$$\vec{y} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \vec{x} = \begin{pmatrix} r \\ \theta \\ z \end{pmatrix}.$$

To solve for r and θ involves inverse trig functions; the answer is not unique, and not smooth on the axis. In particular,

$$\theta = \tan^{-1}\frac{y}{x} + \text{const.},$$

where the constant depends on the quadrant. However, calculating the derivatives of r and θ is easier, as we'll see.

Returning to the general case, let's write the condition defining the inverse function as

$$f\big(f^{-1}(\vec{y})\big) = \vec{y} = I(\vec{y}),$$

where $I\colon \mathbf{R}^n \to \mathbf{R}^n$ is the identity function. (Its matrix is the identity (unit) matrix, 1.) By the chain rule in its Jacobian-matrix form,

$$f'\big(f^{-1}(\vec{y})\big)\ \ {f^{-1}}'(\vec{y}) = I'(\vec{y}) = 1.$$

(Remember that the Jacobian matrix of a linear function is the same as the matrix that represents the function itself.) Therefore,

$${f^{-1}}'(\vec{y}) = \big[f'\big(f^{-1}(\vec{y})\big)\big]^{-1}.$$

Alternative ways of writing this are

$$J_{\vec{y}} f^{-1} = \big[J_{f^{-1}(\vec{y})} f\big]^{-1} \qquad \text{and} \qquad \frac{d\vec{x}}{d\vec{y}} = \left[\left.\frac{d\vec{y}}{d\vec{x}}\right|_{\vec{x}=f^{-1}(\vec{y})}\right]^{-1}.$$

Normally in a calculation the right-hand side of this formula will come out as a function of $\vec{x}$, not of $\vec{y}$.

Example: For the cylindrical coordinates,

$$\frac{d\vec{y}}{d\vec{x}} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \cdots \\ \frac{\partial y}{\partial r} & \cdots & \\ \cdots & & \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Inverting the matrix, we find

$$\frac{d\vec{x}}{d\vec{y}} = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} & \cdots \\ \frac{\partial \theta}{\partial x} & \cdots & \\ \cdots & & \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, for example,

$$\frac{\partial r}{\partial x} = \cos\theta = \frac{x}{r} = \frac{x}{\sqrt{x^2+y^2}}.$$

We can check this directly by inverting the original nonlinear equations and then differentiating; for instance,

$$r = \sqrt{x^2+y^2} \quad \text{leads to} \quad \frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}.$$

Note that for many purposes the expression in terms of polar coordinates ($\cos\theta$) or the mixed expression (x/r) may be as useful as the complicated one in Cartesian coordinates (or more so).

More generally, in any of the many cases where a curvilinear coordinate system is defined by formulas giving the Cartesian coordinates as functions of the curvilinear ones, we have seen in Sec. 4.2 that the columns of the Jacobian matrix of that coordinate transformation are the tangent vectors to the coordinate curves. The discussion above shows that the other set of local basis vectors, the normal vectors to the coordinate surfaces, are the *rows* of the *inverse* of the Jacobian matrix. The formulas for the normal vectors obtained in this way are functions of the curvilinear coordinates, not of the Cartesian ones as they are when we find them by inverting the nonlinear coordinate transformation as in the examples in Sec. 4.2. Let us summarize the discussion in that section and here by a formal theorem:

Theorem: Let $\vec{x} = T(\vec{u})$ be a transformation from curvilinear coordinates to Cartesian coordinates in $\mathbf{R}^n$. The columns of the Jacobian matrix

$$T'(\vec{u}) = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

are tangent vectors to the coordinate curves (not normalized to unit length). More precisely,

$$c_k \equiv \frac{\partial \vec{x}}{\partial u_k} \equiv \left\{ \frac{\partial x_j}{\partial u_k} \right\}\Big|_{j=1,\ldots,n} ,$$

evaluated at $\vec{u}_0$, is tangent to the curve through $\vec{x}_0 \equiv T(\vec{u}_0)$ along which u_k varies and all other components of $\vec{u}$ are fixed. The rows of the inverse Jacobian $T'^{-1} = (T^{-1})'$ are normal vectors to the coordinate surfaces (not normalized to unit length). (These "surfaces" have dimension $n-1$.) More precisely,

$$\vec{d}_k \equiv \nabla u_k \equiv \left(\frac{\partial u_k}{\partial x_1} , \frac{\partial u_k}{\partial x_2} , \ldots \right) ,$$

evaluated at $\vec{u}_0$, is normal (perpendicular) to the surface through $\vec{x}_0 \equiv T(\vec{u}_0)$ on which u_k is constant and on which all other components of $\vec{u}$ vary and serve as coordinates.

So far, we have *assumed* that f^{-1} exists and found a formula for its differential (or its Jacobian matrix) and hence its partial derivatives. (In

doing so, we also tacitly *assumed* that the Jacobian matrix is invertible.) The full truth is even better:

Inverse Function Theorem: If $\vec{f}\colon \mathbf{R}^n \to \mathbf{R}^n$ is differentiable, if the matrix $\vec{f}\,'(\vec{x}_0) \equiv \left\{\frac{\partial y_j}{\partial x_k}\right\}\Big|_{\vec{x}_0}$ is invertible (equivalently, if the linear transformation $d_{\vec{x}_0}\vec{f}$ has maximal rank), and if the first partial derivatives of $\vec{f}$ (the matrix elements of $\vec{f}\,'(\vec{x})$) are continuous for $\vec{x}$ near $\vec{x}_0$, then there is a region $D \subseteq \mathbf{R}^n$ around $\vec{x}_0$ which is the range of a local inverse function, $\vec{f}^{\,-1}$, whose domain is the image set $\vec{f}[D]$ and whose derivative (Jacobian) matrix is $\vec{f}\,'(\vec{x})^{-1}$.

A full proof of this theorem is beyond the scope of this course. Proofs can be found in textbooks on rigorous advanced calculus or real analysis, such as L. H. Loomis and S. Sternberg, *Advanced Calculus* (Addison–Wesley, 1968), and W. Rudin, *Principle of Mathematical Analysis*, 3rd ed. (McGraw–Hill, 1976). What is hard to prove is that the inverse function *exists*; once that is known, it follows from our calculation above that the derivative is as stated.

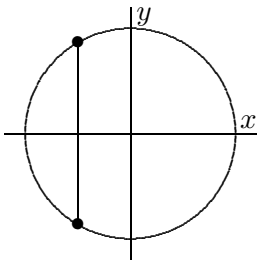
The implicit function theorem

Differentiation of inverse functions is a special case of differentiation of *implicit functions*. Recall how that goes in $\mathbf{R}^1$:

$$x^2 + y^2 = 25$$

$$\text{implies } 2x + 2y\,\frac{dy}{dx} = 0;$$

$$\text{therefore, } \frac{dy}{dx} = -\,\frac{x}{y}\,.$$



Which function $y = f(x)$ you are differentiating depends on which value of y you plug in (the positive or the negative square root).

In higher dimensions, consider an equation

$$\vec{G}(\vec{x},\vec{y}) = \vec{0}, \qquad \text{where} \quad \vec{G}\colon \mathbf{R}^{n+m} \to \mathbf{R}^m, \quad \vec{x} \in \mathbf{R}^n, \quad \vec{y} \in \mathbf{R}^m.$$

The condition may locally determine $\vec{y}$, given $\vec{x}$. (It amounts to m equations in m unknowns, possibly nonlinear equations.) Let's assume so. That is,

locally we'll have a function

$$\vec{y} = \vec{f}(\vec{x}), \qquad \vec{f}\colon \mathbf{R}^n \to \mathbf{R}^m.$$

(Examples are coming up soon.)

To proceed further, we need to invent a notation adequate for the occasion. Let $\partial\vec{G}/\partial\vec{x}$ stand for the matrix of partial derivatives of the components of $\vec{G}$ with respect to the $\vec{x}$ variables (with $\vec{y}$ fixed), and so forth. That is,

$$\frac{d\vec{G}}{d(\vec{x},\vec{y})} \equiv \left(\begin{array}{ccc|ccc} \frac{\partial G_1}{\partial x_1} & \cdots & \frac{\partial G_1}{\partial x_n} & \frac{\partial G_1}{\partial y_1} & \cdots & \\ \vdots & & & & & \\ \frac{\partial G_m}{\partial x_1} & \cdots & & \cdots & & \frac{\partial G_m}{\partial y_m} \end{array}\right) = \left(\begin{array}{c|c} \dfrac{\partial\vec{G}}{\partial\vec{x}} & \dfrac{\partial\vec{G}}{\partial\vec{y}} \end{array}\right).$$

Then set $\vec{y} = \vec{f}(\vec{x})$ in the condition $\vec{G}(\vec{x},\vec{y}) = 0$ and differentiate the latter with respect to $\vec{x}$, invoking the chain rule:

$$0 = \frac{d\vec{G}\big(\vec{x},\vec{f}(\vec{x})\big)}{d\vec{x}} = \frac{\partial\vec{G}}{\partial\vec{x}} + \frac{\partial\vec{G}}{\partial\vec{y}}\,\frac{\partial\vec{y}}{\partial\vec{x}}\,.$$

In more detail:

$$\frac{d\vec{G}}{d\vec{x}} = \frac{d\vec{G}}{d(\vec{x},\vec{y})}\,\frac{d\big(\vec{x},f(\vec{x})\big)}{d\vec{x}}\,, \qquad \frac{d\big(\vec{x},f(\vec{x})\big)}{d\vec{x}} = \left(\begin{array}{ccc} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \vdots & \vdots & \ddots \\ \hline & \dfrac{d\vec{y}}{d\vec{x}} & \end{array}\right).$$

Therefore,

$$\frac{d\vec{y}}{d\vec{x}} = -\left(\frac{\partial\vec{G}}{\partial\vec{y}}\right)^{-1}\frac{\partial\vec{G}}{\partial\vec{x}}\,. \qquad (*)$$

This is the formula we are looking for.

Of course, the formula $(*)$ makes no sense unless the (square) matrix $\partial\vec{G}/\partial\vec{y}$ has maximal rank. It seems plausible that this condition is closely related to our initial assumption that the equations $\vec{G}(\vec{x},\vec{y})$ can be solved (in principle) for $\vec{y}$. Indeed, just as in the inverse-function case, the full implicit function theorem guarantees existence of the implicit function, given a rank condition on the matrix of derivatives:

Implicit Function Theorem: Let $(\vec{x}_0, \vec{y}_0)$ be a solution of the equation $\vec{G}(\vec{x}, \vec{y}) = 0$. If

$$\left(\frac{\partial \vec{G}}{\partial \vec{y}} \right)^{-1} \Bigg|_{(\vec{x}_0, \vec{y}_0)}$$

exists and $\vec{G}$ is continuously differentiable in the vicinity, then a local implicit function, $\vec{y} = \vec{f}(\vec{x})$, such that $\vec{G}\big(\vec{x}, f(\vec{x})\big) = 0$, exists near $\vec{x}_0$, and its Jacobian matrix is given by $(*)$.

Exactly the same remarks about proof apply here as to the inverse function theorem.

In practice, you may not do such a calculation in matrix notation; as in the case of the chain rule, you can let the Leibnitz notation take care of the matrix multiplication for you automatically. This "classical" approach is to differentiate each component equation of $\vec{G}(\vec{x}, \vec{y}) = 0$ with respect to each component of $\vec{x}$ and then solve for all the $\frac{\partial y_j}{\partial x_k}$.

Example 1. Regard the equations

$$\begin{aligned} x^2 - y^2 + u^2 - v^2 &= 0 \\ xy - uv &= 0 \end{aligned}$$

as implicitly defining u and v as functions of x and y in a neighborhood of the solution

$$x = 1, \qquad y = 1, \qquad u = 1, \qquad v = 1.$$

Let's calculate $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial x}$ at that point. Differentiate the equations with respect to x (remembering that u and v are functions of x, but y is independent):

$$\begin{aligned} 2x + 2u \frac{\partial u}{\partial x} - 2v \frac{\partial v}{\partial x} &= 0, \\ y - \frac{\partial u}{\partial x} v - u \frac{\partial v}{\partial x} &= 0. \end{aligned}$$

Evaluate the coefficients at the point:

$$\begin{aligned} \frac{\partial u}{\partial x} - \frac{\partial v}{\partial x} &= -1, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x} &= 1. \end{aligned}$$

The solution is

$$\frac{\partial u}{\partial x} = 0, \qquad \frac{\partial v}{\partial x} = 1.$$

We could calculate the derivatives with respect to y similarly. The two calculations fit together to make up the matrix calculation prescribed in $(*)$ (with $\vec{y} = (u, v)$, $\vec{x} = (x, y)$).

Example 2. Three quantities A, B, C are related by the equations

$$\begin{aligned} AB + BC &= 18, \\ A + B &= 2C. \end{aligned}$$

(a) Regard A and B as functions of C. Find the derivatives of A and B with respect to C at the point where $A = B = C = 3$. (b) Find the derivatives of functions $f(A, B)$ and $g(A, B)$ with respect to C at that point, if

$$\begin{pmatrix} \frac{\partial f}{\partial A} & \frac{\partial f}{\partial B} \\ \frac{\partial g}{\partial A} & \frac{\partial g}{\partial B} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

there.

SOLUTION: Implicit differentiation with respect to C yields

$$\begin{aligned} \frac{dA}{dC}\, B + A\, \frac{dB}{dC} + \frac{dB}{dC}\, C + B &= 0, \\ \frac{dA}{dC} + \frac{dB}{dC} &= 2. \end{aligned}$$

Setting all variables equal to 3 (which, we note in passing, does satisfy the original equations!), we get

$$\begin{aligned} 3\frac{dA}{dC} + 6\, \frac{dB}{dC} &= -3, \\ \frac{dA}{dC} + \frac{dB}{dC} &= 2. \end{aligned}$$

The solution (by elementary algebra or by inverting the coefficient matrix) is

$$\begin{pmatrix} \frac{dA}{dC} \\ \frac{dB}{dC} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & 2 \\ \frac{1}{3} & -1 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}.$$

This completes part (a). Part (b) can now be done by the chain rule:

$$\begin{pmatrix} \frac{df}{dC} \\ \frac{dg}{dC} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial A} & \frac{\partial f}{\partial B} \\ \frac{\partial g}{\partial A} & \frac{\partial g}{\partial B} \end{pmatrix} \begin{pmatrix} \frac{dA}{dC} \\ \frac{dB}{dC} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ -3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Next, a larger-scale example of implicit differentiation and of nonorthogonal curvilinear coordinates:

Example 3. Recall the formulas defining spherical coordinates:

$$x = r\sin\theta\cos\phi, \tag{1}$$

$$y = r\sin\theta\sin\phi, \tag{2}$$

$$z = r\cos\theta. \tag{3}$$

These are 3 equations in 6 variables. Therefore, at most points in the space we should be able to solve for 3 of the variables in terms of the other 3. To consider (r, θ, ϕ) as functions of (x, y, z) is a routine application of the *inverse* function theorem. (Some details will be given as an example in Sec. 7.1.) In order to see an application of the more general *implicit* function theorem, let's consider (x, θ, ϕ) as functions of (r, y, z). More generally, we shall study the possibility of using (r, y, z) as a coordinate system for $\mathbf{R}^3$, or for parts of it.

(1) First we ask the standard calculational question: What are the partial derivatives of the dependent variables with respect to the independent ones (the new coordinates)? That is, what are the elements of the Jacobian matrix

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial y} & \frac{\partial x}{\partial z} \\ \frac{\partial \theta}{\partial r} & \frac{\partial \theta}{\partial y} & \frac{\partial \theta}{\partial z} \\ \frac{\partial \phi}{\partial r} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{pmatrix}? \tag{4}$$

To find the first column of this matrix, differentiate (1)–(3) with respect to r, with y and z fixed:

$$\frac{\partial x}{\partial r} = \sin\theta\cos\phi + r\cos\theta\cos\phi\,\frac{\partial\theta}{\partial r} - r\sin\theta\sin\phi\,\frac{\partial\phi}{\partial r}\,, \tag{5}$$

$$0 = \frac{\partial y}{\partial r} = \sin\theta\sin\phi + r\cos\theta\sin\phi\,\frac{\partial\theta}{\partial r} + r\sin\theta\cos\phi\,\frac{\partial\phi}{\partial r}\,, \tag{6}$$

$$0 = \frac{\partial z}{\partial r} = \cos\theta - r\sin\theta\,\frac{\partial\theta}{\partial r}\,. \tag{7}$$

As it happens, this is an essentially triangular system that can be solved from the bottom up: We easily get

$$\frac{\partial\theta}{\partial r} = \frac{1}{r\tan\theta}\,, \tag{8}$$

$$\frac{\partial \phi}{\partial r} = -\frac{\tan \phi}{r \sin^2 \theta}\,, \tag{9}$$

and a messy expression that simplifies to*

$$\frac{\partial x}{\partial r} = \csc \theta \sec \phi. \tag{10}$$

(2) Let us pause to examine where the three formulas (8)–(10) break down. We observe that their denominators can vanish under any of three conditions:

(i) $r = 0$,

(ii) $\sin \theta = 0$,

(iii) $\cos \phi = 0$.

Cases (i) and (ii) are no surprises: The origin is always a singular point in spherical coordinates, and the north and south polar axes ($\theta = 0$ or π) only slightly less singular. The third condition is less obvious. We observe (from (1)) that wherever (i) and (ii) are false, (iii) is equivalent to $x = 0$. What is special about that plane, with regard to the coordinate system (r, y, z)? Consider what happens when we hold the coordinates y and z fixed and vary r. The resulting "curve" is a horizontal line, and varying r along it is equivalent to varying x (or ϕ). As x varies from $-\infty$ to $+\infty$, r decreases until $x = 0$, then begins to increase again. As the line crosses the plane, the line becomes *tangent* to one of the spheres of constant r, rather than cutting through it. At this point, a tiny change in r corresponds to a large change in x and in ϕ, and this is exactly what the derivatives (8) and (10) are telling us.

---

* Incidentally, a first attempt to use a computer algebra program to solve the entire linear system (5)–(7) yielded

$$\frac{\partial x}{\partial r} = \cos \phi \sin \theta - \frac{\cot \theta \cos \theta \sec \phi - r^2 \cos \theta \sin \theta}{r^2 + \sin \phi \sin \theta \tan \phi}\,.$$

This is *obviously wrong*, since the "units" don't make sense — the square of a length shouldn't be added to a pure trigonometric function. In fact, it is easy to see from (5), (8), and (9) that the correct answer cannot depend upon r at all. This anecdote demonstrates the importance of maintaining skepticism about the output of computations, especially symbolic ones, and of developing the gut understanding of mathematical and physical concepts that enables one to catch errors in calculations, whether by machine or by hand.

(3) To investigate this phenomenon more systematically, we recall that the implicit function theorem guarantees existence and differentiability of an implicit function wherever a certain determinant is nonzero. This is the determinant of the Jacobian matrix of the (3-component) function of (x, θ, ϕ) obtained by moving all the terms in (1)–(3) to the same side and treating all 6 variables as independent. (This is the matrix called $\frac{\partial \vec{G}}{\partial \vec{y}}$ above.) We have already calculated these partial derivatives in (5)–(7); we merely need to extract them with the correct signs and in the correct positions:

$$\begin{vmatrix} -1 & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ 0 & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ 0 & -r\sin\theta & 0 \end{vmatrix} = -r^2\cos\phi\sin^2\theta. \tag{11}$$

The zeros of the determinant are exactly the same "bad spots" we identified earlier. They mark the places where the quantities (x, θ, ϕ) do not depend smoothly on (r, y, z).

We can complete the calculation of the partial derivatives (4) by multiplying the negative of the inverse of the matrix in (11), which is

$$\begin{pmatrix} 1 & \tan\phi & \cot\theta\sec\phi \\ 0 & 0 & \frac{1}{r}\csc\theta \\ 0 & -\frac{1}{r}\csc\theta\sec\phi & -\frac{1}{r}\cot\theta\csc\theta\tan\phi \end{pmatrix}, \tag{12}$$

by the matrix of derivatives of the 6-variable function with respect to the other set of variables, (r, y, z) (the matrix called $\frac{\partial \vec{G}}{\partial \vec{x}}$ above), which is

$$\begin{pmatrix} \sin\theta\cos\phi & 0 & 0 \\ \sin\theta\sin\phi & -1 & 0 \\ \cos\theta & 0 & -1 \end{pmatrix}. \tag{13}$$

The result is

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial y} & \frac{\partial x}{\partial z} \\ \frac{\partial \theta}{\partial r} & \frac{\partial \theta}{\partial y} & \frac{\partial \theta}{\partial z} \\ \frac{\partial \phi}{\partial r} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{pmatrix} = \begin{pmatrix} \csc\theta\sec\phi & -\tan\phi & -\cot\theta\sec\phi \\ \frac{1}{r}\cot\phi & 0 & -\frac{1}{r}\csc\theta \\ -\frac{1}{r}\csc^2\theta\tan\phi & \frac{1}{r}\csc\theta\sec\phi & \frac{1}{r}\cot\theta\csc\theta\tan\phi \end{pmatrix}. \tag{14}$$

(4) We saw above that the lack of smoothness (and possible nonexistence) of the implicit function in this example is at least as much the fault of the behavior of the independent variables as of the dependent ones. This raises the question: In what regions is (r, y, z) a well-defined, smooth coordinate system? We can answer this question by looking at the local bases

defined (where they exist) by the tangent vectors to the coordinate curves and the normal vectors to the coordinate surfaces. These fit together to form the Jacobian matrices of the function $(r, y, z) \mapsto (x, y, z)$ and its inverse, so in essence this is another calculation of the same type as the foregoing.

It is convenient to use the explicit formula

$$r^2 = x^2 + y^2 + z^2. \tag{15}$$

From this we get a system of formulas analogous to (1)–(3):

$$x = \sqrt{r^2 - y^2 - z^2}, \tag{16}$$

$$y = y, \tag{17}$$

$$z = z. \tag{18}$$

Derivatives of x with respect to the 3 independent variables can be calculated either from (16) or (slightly less messily) by implicit differentiation of (15). We find that the tangent vectors are

$$\frac{\partial \vec{x}}{\partial r} = \begin{pmatrix} \frac{r}{x} \\ 0 \\ 0 \end{pmatrix}, \qquad \frac{\partial \vec{x}}{\partial y} = \begin{pmatrix} -\frac{y}{x} \\ 1 \\ 0 \end{pmatrix}, \qquad \frac{\partial \vec{x}}{\partial z} = \begin{pmatrix} -\frac{z}{x} \\ 0 \\ 1 \end{pmatrix}. \tag{19}$$

The normal vectors can be obtained directly from (15), (17), and (18):

$$\nabla r = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right), \qquad \nabla y = (0, 1, 0), \qquad \nabla z = (0, 0, 1). \tag{20}$$

It is easy to check that the matrices formed from the vectors in (19) and (20) are inverse to each other, as they should be.

Let us examine (19) and (20) for singularities:

(i) When $r = 0$ (and hence $x = y = z = 0$) most of the vectors are not defined. (The expressions involve division of 0 by 0. Going back to the definitions of the vectors as derivatives, one sees that those derivatives truly don't exist.)

(ii) The polar axis is *not* singular in this problem.

(iii) When $x = 0$ but $r \neq 0$, ∇r is a linear combination of ∇y and ∇z. Algebraically, this says that the Jacobian determinant of the coordinate transformation vanishes. Geometrically, it says that

the "r direction" lies in the y–z plane, so that there is no coordinate to measure displacements transverse to the plane. The coordinate system breaks down there. Since the determinant is 0, the determinant of the inverse transformation should in some sense be infinite; indeed, we see that each of the tangent-vector formulas involves division of a nonzero quantity by zero, and the Jacobian matrix of the inverse transformation does not exist at points on the plane. All these facts are expressions of the phenomenon we noted earlier: The surface of constant r becomes tangent to the curves of constant y and z.

Postscript

The linear algebra sections of Chapters 3, 4, and 5 have unavoidably presented a melange of abstract concepts at a rate that makes many students uncomfortable. Furthermore, as previously remarked, there is no ideal order in which to learn these things; they have to be understood as a whole. This is an excellent time to pause and consolidate your understanding by rereading the definitions and theorems in Sections 3.1, 3.2, 4.1, 4.3, 4.4, 4.5, 5.1, 5.2, 5.4.

Exercises

5.5.1 Verify that the inverse function theorem is a special case of the implicit function theorem. (How are the G, $\vec{x}$, and $\vec{y}$ of the latter related to the f, $\vec{x}$, and $\vec{y}$ of the former?)

5.5.2 Let $x = u^2 - \cos v$, $y = e^u + v^2$. Calculate $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial x}$, $\frac{\partial u}{\partial y}$, and $\frac{\partial v}{\partial y}$ as functions of u and v.

5.5.3 A function $\vec{u} = F(\vec{v})$ is defined implicitly by the equations

$$\begin{aligned} u_1 - {u_2}^2 + 3u_1v_2 &= 0, \\ {u_1}^2v_2 + u_2{v_1}^2 &= 0. \end{aligned}$$

Find the Jacobian matrix, F', (at points where it exists) by implicit differentiation.

5.5.4 *(Continuation of Exercise 3.4.3)*

(a) Find the partial derivatives $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial x}$ as functions of u and v. Do this by the "classical" method: Differentiate each of the two formulas

$$x = \cosh u \cos v, \qquad y = \sinh u \sin v$$

with respect to x, then solve for the desired derivatives. (Find formulas for the partials at a *general point*, not just at $\vec{u}_0$.)

(b) Let $\vec{x}_0 = F(\vec{u}_0)$ ($\vec{u}_0 \equiv \binom{1}{0}$). From your answer to 3.4.3(a), find the matrix representing $d_{\vec{x}_0}F^{-1}$. Verify the consistency of your result with (a) of the present problem.

5.5.5 Find the formulas for $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial x}$, $\frac{\partial w}{\partial x}$ in *parabolic cylindrical coordinates*,

$$x = 2uv, \qquad y = u^2 - v^2, \qquad z = w.$$

(The answers will be functions of (u, v, w).)

5.5.6 A new coordinate system (u, v) is introduced into the plane by

$$x = u^3 - 3uv^2,$$
$$y = 3u^2v - v^3.$$

(a) Find (as functions of u and v) the Cartesian components of the (unnormalized) tangent vectors to the coordinate curves (the curves $v =$ constant and the curves $u =$ constant).

(b) Calculate the determinant of the Jacobian matrix.

(c) Find the normal vectors to the coordinate surfaces (as functions of u and v).

5.5.7 Define coordinates q and p in a region of the (y, z) plane by

$$y = q^2 \sinh p, \qquad z = q^2 \cosh p.$$

Find formulas for $\frac{\partial q}{\partial y}$, $\frac{\partial q}{\partial z}$, $\frac{\partial p}{\partial y}$, and $\frac{\partial p}{\partial z}$ by implicit differentiation. (Answers will be functions of q and p.)

5.5.8 Define coordinates (u, v) in a suitable region of the (x, y) plane by

$$x = u^2 v, \qquad y = u + v^2.$$

(a) To be "suitable", the region must stay away from the places where the determinant of the Jacobian of the coordinate transformation equals 0. Find these dangerous places. (The final answer should be one or more equations relating x and y, with u and v eliminated.)

(b) Calculate the partial derivatives of u and v with respect to x and y. (The formulas for the answers will contain u and v.)

(c) In the (x, y) plane, at the point $(4, 3)$ (corresponding to $u = 2$, $v = 1$), draw the tangent vectors to the coordinate curves and the normal vectors to the coordinate "surfaces" (which are curves, in dimension 2). (Don't normalize these vectors to unit length.) Then plot the curve $u = 2$, the curve $v = 1$, and (with dashed lines) the dangerous curves you found in (a). Comment on the relation among these curves.

5.5.9 Suppose that $x^2 y + yz = 0$ and $xyz + 1 = 0$.

(a) Find $\frac{dx}{dy}$ and $\frac{dz}{dy}$ at $(x, y, z) = (1, 1, -1)$.

(b) What is the function G in the implicit function theorem as applied to this problem?

5.5.10 Again suppose that $x^2 y + yz = 0$ and $xyz + 1 = 0$, but this time regard x and y as functions of z. Find the equation of the tangent line to the curve at any point where $z = -2$. How many such points are there?

5.5.11 Find an equation for the tangent plane to the graph of a function $z = f(x, y)$ at (x_0, y_0) by first finding a normal vector to the surface $G(x, y, z) \equiv f(x, y) - z = 0$. Show that the result agrees with the construction of the tangent plane in Sec. 2.4.

5.5.12 Which statement is correct? In a curvilinear coordinate system, (u, v, w), in $\mathbf{R}^3$, the normal vector to the coordinate surface $v =$ constant (at a particular point $\vec{x}$) is

(A) always perpendicular to the normal vector (at $\vec{x}$) to the surface $u =$ constant.

(B) always perpendicular to the tangent vectors (at $\vec{x}$) to the surface $v = \text{const}$.

(C) always of length 1 when computed from the Jacobian matrix.

(D) equal to the gradient of either u or w.

5.5.13 The quantities A, B, s, and t are related by the equations

$$\begin{aligned} tA^4 - sB^4 &= 0, \\ sA^2 + tB^2 &= 0. \end{aligned}$$

Find formulas for $\frac{\partial A}{\partial t}$ and $\frac{\partial B}{\partial t}$ (partial derivatives taken with s fixed). These formulas will involve all 4 variables.

5.5.14 *(Continuation of Exercise 4.2.1)* For elliptic coordinates,

(a) Find all four partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$.

(b) Find the normal vectors to the coordinate curves.

5.5.15 *(Continuation of Exercise 4.2.3)* Find the normal vectors to the coordinate surfaces for spherical coordinates in $\mathbf{R}^3$.

5.5.16 Use implicit differentiation to solve Exercise 4.2.5 in another way. (This time the answers will be functions of u and v, whereas before they were functions of x and y.)