5.4 Independence, Span and Basis

The technical topics of independence, dependence and span apply to the study of Euclidean spaces $\mathcal{R}^2$, $\mathcal{R}^3$, $\ldots$, $\mathcal{R}^n$ and also to the continuous function space $C(E)$, the space of differentiable functions $C^1(E)$ and its generalization $C^n(E)$, and to general abstract vector spaces.

Basis and General Solution

The term **basis** has been introduced earlier for systems of linear algebraic equations. To review, a basis is obtained from the vector general solution $\vec{x}$ of matrix equation $A\vec{x} = \vec{0}$ by computing the partial derivatives ∂_{t_1}, ∂_{t_2}, ... of $\vec{x}$, where t_1, t_2, ... is the list of invented symbols assigned to the free variables identified in $\mathbf{rref}(A)$.

The partial derivatives are **special solutions** to the homogeneous equation $A\vec{x} = \vec{0}$. Knowing the special solutions is sufficient for writing out the general solution. In summary:

> A **basis** is an abbreviation or shortcut notation for the general solution.

Deeper properties have been isolated for the list of special solutions obtained from the partial derivatives $\partial_{t_1}\vec{x}$, $\partial_{t_2}\vec{x}$, ... The most important properties are **span** and **independence**.

Independence, Span and Basis

A list of vectors $\vec{v}_1$, ..., $\vec{v}_k$ is said to **span** a vector space V (definition on page 297), written

$$V = \mathbf{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k),$$

provided V consists of exactly the set of all linear combinations

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k.$$

The notion originates with the general solution $\vec{v}$ of a homogeneous matrix system $A\vec{v} = \vec{0}$, where the invented symbols t_1, ..., t_k are the constants c_1, ..., c_k and the vector partial derivative list $\partial_{t_1}\vec{v}$, ..., $\partial_{t_k}\vec{v}$ is the list of special solution vectors $\vec{v}_1$, ..., $\vec{v}_k$.

Vectors $\vec{v}_1$, ..., $\vec{v}_k$ are said to be **independent** provided each linear combination $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is represented by a unique set of constants c_1, ..., c_k. See pages 369 and 375 for independence tests.

Definition 6 (Basis)
A **basis** of a vector space V is defined to be a list of independent vectors $\vec{v}_1, \ldots, \vec{v}_k$ which spans V. A basis is tested by two checkpoints:

1. The list of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ is independent.

2. The vectors span V, written $V = \mathbf{span}(\vec{v}_1, \ldots, \vec{v}_k)$.

Bases are used to express the **general solution** of a linear problem. The **spanning condition** means that every possible vector in V is a linear combination of basis elements. The **independence condition** means that linear combinations are uniquely represented, which, in practical terms, means that the general solution expression has *the fewest possible terms.*

The Vector Spaces $\mathcal{R}^n$

The vector space $\mathcal{R}^n$ of n-element fixed column vectors (or row vectors) is from the view of applications a *storage system for organization of numerical data sets* that is equipped with an algebraic toolkit. The organizational scheme induces a *data structure* onto the numerical data set. In particular, whether needed or not, there are pre-defined operations of addition (+) and scalar multiplication ($\cdot$) which apply to fixed vectors. The two operations on fixed vectors satisfy the *closure law* and in addition obey the *eight algebraic vector space properties.* We view the vector space $V = \mathcal{R}^n$ as the **data set** consisting of data item packages. The **toolkit** is the following set of eight algebraic properties.

Closure	The operations $\vec{X}+\vec{Y}$ and $k\vec{X}$ are defined and result in a new vector which is also in the set V.	
Addition	$\vec{X}+\vec{Y}=\vec{Y}+\vec{X}$	commutative
	$\vec{X}+(\vec{Y}+\vec{Z})=(\vec{Y}+\vec{X})+\vec{Z}$	associative
	Vector $\vec{0}$ is defined and $\vec{0}+\vec{X}=\vec{X}$	zero
	Vector $-\vec{X}$ is defined and $\vec{X}+(-\vec{X})=\vec{0}$	negative
Scalar multiply	$k(\vec{X}+\vec{Y})=k\vec{X}+k\vec{Y}$	distributive I
	$(k_1+k_2)\vec{X}=k_1\vec{X}+k_2\vec{X}$	distributive II
	$k_1(k_2\vec{X})=(k_1k_2)\vec{X}$	distributive III
	$1\vec{X}=\vec{X}$	identity

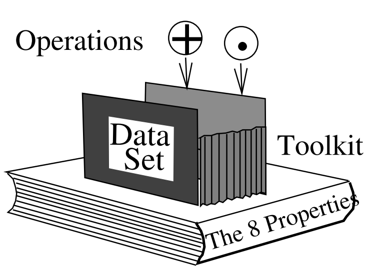

Figure 12. A Data Storage System.
A vector space is a data set of data item packages plus a storage system which organizes the data. A toolkit is provided consisting of operations + and $\cdot$ plus 8 algebraic vector space properties.

Fixed Vectors and the Toolkit

Scalar multiplication of fixed vectors is commonly used for re-scaling, especially to unit systems fps, cgs and mks. For instance, a numerical data set of lengths recorded in meters (mks) is re-scaled to centimeters (cgs) using scale factor $k = 100$.

Addition and subtraction of fixed vectors is used in a variety of calculations, which includes averages, difference quotients and calculus operations like integration.

Planar Plots and the Toolkit

The data set for a plot problem consists of the plot points in $\mathcal{R}^2$, which are the **dots** for the connect-the-dots graphic. Assume the function $y(x)$ to be plotted comes from a differential equation like $y' = f(x, y)$, then Euler's numerical method could be used for the sequence of dots in the graphic. In this case, the next dot is represented as $\vec{v}_2 = \vec{v}_1 + \vec{E}(\vec{v}_1)$. Symbol $\vec{v}_1$ is the previous dot and symbol $\vec{E}(\vec{v}_1)$ is the Euler increment. We define

$$\vec{v}_1 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad \vec{E}(\vec{v}_1) = h \begin{pmatrix} 1 \\ f(x_0, y_0) \end{pmatrix},$$
$$\vec{v}_2 = \vec{v}_1 + \vec{E}(\vec{v}_1) = \begin{pmatrix} x_0 + h \\ y_0 + hf(x_0, y_0) \end{pmatrix}.$$

A step size $h = 0.05$ is commonly used. The Euler increment $\vec{E}(\vec{v}_1)$ is given as scalar multiplication by h against an $\mathcal{R}^2$-vector which involves evaluation of f at the previous dot $\vec{v}_1$.

In summary, the **dots** for the graphic of $y(x)$ form a data set in the vector space $\mathcal{R}^2$. The dots are obtained by algorithm rules, which are easily expressed by vector addition (+) and scalar multiplication ($\cdot$). The 8 properties of the toolkit were used in a limited way.

Digital Photographs

A digital photo consists of many **pixels** arranged in a two dimensional array. Structure can be assigned to the photo by storing the pixel digital color data in a matrix A of size $n \times m$. Each entry of A is an integer which encodes the color information at a specific pixel location.

The set V of all $n \times m$ matrices is a vector space under the usual rules for matrix addition and scalar multiplication. Initially, V is just a storage system for photos. However, the algebraic toolkit for V is a convenient way to express operations on photos. We give one illustration: breaking a photo into RGB (Red, Green, Blue) separation photos, in order to make color separation transparencies.

One way to encode each entry of A is to define $a_{ij} = r_{ij} + g_{ij}x + b_{ij}x^2$ where x is some convenient base. The integers r_{ij}, g_{ij}, b_{ij} represent the amount of red, green and blue present in the pixel with data a_{ij}. Then $A = R + Gx + Bx^2$ where $R = [r_{ij}]$, $G = [g_{ij}]$, $B = [b_{ij}]$ are $n \times m$ matrices that represent the color separation photos. These monochromatic photos are superimposed as color transparencies on a standard overhead projector to duplicate the original photograph.

Printing machinery from many years ago employed separation negatives and multiple printing runs in primary ink colors to make book photos. The advent of digital printers and simpler inexpensive technologies has made the separation process nearly obsolete. To help the reader understand the historical events, we record the following quote from Sam Wang[7]:

> I encountered many difficulties when I first began making gum prints: it was not clear which paper to use; my exposing light (a sun lamp) was highly inadequate; plus a myriad of other problems. I was also using panchromatic film, making in–camera separations, holding RGB filters in front of the camera lens for three exposures onto 3 separate pieces of black and white film. I also made color separation negatives from color transparencies by enlarging in the darkroom. Both of these methods were not only tedious but often produced negatives very difficult to print — densities and contrasts that were hard to control and working in the dark with panchromatic film was definitely not fun. The fact that I got a few halfway decent prints is something of a small miracle, and represents hundreds of hours of frustrating work! Digital negatives by comparison greatly simplify the process. Nowadays (2004) I use color images from digital cameras as well as scans from slides, and the negatives print much more predictably.

Function Spaces

The default storage system used for applications involving ordinary or partial differential equations is a *function space*. The data item packages for differential equations are their solutions, which are *functions*, or in an applied context, a graphic defined on a certain graph window. They are **not** column vectors of numbers.

Functions and Column Vectors

An alternative view, adopted by researchers in numerical solutions of differential equations, is that a solution is a table of numbers, consisting of pairs of x and y values.

[7]Sam Wang teaches photography and art with computer at Clemson University in South Carolina. His photography degree is from the University of Iowa (1966). **Reference**: *A Gallery of Tri-Color Prints*, by Sam Wang

These researchers might view a function as being a fixed vector. Their unique intuitive viewpoint is that a function is a **graph** and a graph is determined by so many **dots**, which are practically obtained by **sampling** the function $y(x)$ at a reasonably dense set of x-values. The approximation is

$$y \approx \begin{pmatrix} y(x_1) \\ y(x_2) \\ \vdots \\ y(x_n) \end{pmatrix}$$

where $x_1, \ldots, x_n$ are the **samples** and $y(x_1), \ldots, y(x_n)$ are the **sampled values** of function y.

The trouble with the approximation is that two different functions may need different sampling rates to properly represent their graphic. The result is that the two functions might need data storage systems of different dimensions, e.g., f needs its sampled values in $\mathcal{R}^{200}$ and g needs its sampled values in $\mathcal{R}^{400}$. The absence of a universal fixed vector storage system for sampled functions explains the appeal of a storage system like the set of all functions.

Infinitely Long Column Vectors. Novices suggest a way around the lack of a universal numerical data storage system for sampled functions: *develop a theory of column vectors with infinitely many components.* It may help you to think of any function f as an infinitely long column vector, with one entry $f(x)$ for each possible sample x, e.g.,

$$\vec{f} = \begin{pmatrix} \vdots \\ f(x) \\ \vdots \end{pmatrix} \qquad \text{level } x$$

It is not clear how to order or address the entries of such a column vector: at algebraic stages it hinders. Can computers store infinitely long column vectors? The easiest path through the algebra is to deal exactly with functions and function notation. Still, there is something attractive about the change from sampled approximations to a single column vector with infinite extent:

$$\vec{f} \approx \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} \to \begin{pmatrix} \vdots \\ f(x) \\ \vdots \end{pmatrix} \qquad \text{level } x$$

The thinking behind the *level* x annotation is that x stands for one of the infinite possibilities for an invented sample. Alternatively, with a rich set of invented samples $x_1, \ldots, x_n$, value $f(x)$ equals approximately $f(x_j)$, where x is closest to some sample x_j.

The Vector Space V of all Functions on a Set E

The rules for function addition and scalar multiplication come from college algebra and pre-calculus backgrounds:

$$(f+g)(x) = f(x) + g(x), \quad (cf)(x) = c \cdot f(x).$$

These rules can be motivated and remembered by the level x notation of infinitely long column vectors:

$$c_1\vec{f} + c_2\vec{g} = c_1 \begin{pmatrix} \vdots \\ f(x) \\ \vdots \end{pmatrix} + c_2 \begin{pmatrix} \vdots \\ g(x) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ c_1 f(x) + c_2 g(x) \\ \vdots \end{pmatrix}$$

The rules define **addition** and **scalar multiplication** of functions. The closure law for a vector space holds. Routine, but tedious, justifications show that V, under the above rules for addition and scalar multiplication, has the required 8-property toolkit to make it a vector space:

Closure	The operations $f+g$ and kf are defined and result in a new function which is also in the set V of all functions on the set E.	
Addition	$f+g=g+f$	commutative
	$f+(g+h)=(f+g)+h$	associative
	The zero function 0 is defined and $0+f=f$	zero
	The function $-f$ is defined and $f+(-f)=0$	negative
Scalar multiply	$k(f+g)=kf+kg$	distributive I
	$(k_1+k_2)f=k_1f+k_2f$	distributive II
	$k_1(k_2f)=(k_1k_2)f$	distributive III
	$1f=f$	identity

Important subspaces of the vector space V of all functions appear in applied literature as the storage systems for solutions to differential equations and solutions of related models.

The Space $C(E)$

Let $E = \{x \;:\; a < x < b\}$ be an open interval on the real line. The set $C(E)$ is defined to be the subset S of the set V of all functions on E obtained by restricting the function to be continuous. Because sums and scalar multiples of continuous functions are continuous, then $S = C(E)$ is a subspace of V and a vector space in its own right. What has been said for an open interval E holds also for an open bounded set E.

The Space $C^1(E)$

The set $C^1(E)$ is the subset of the vector space $C(E)$ of all continuous functions on E obtained by restricting the function to be continuously differentiable. Because sums and scalar multiples of continuously differentiable functions are continuously differentiable, then $C^1(E)$ is a subspace of $C(E)$ and a vector space in its own right.

The Space $C^k(E)$

The set $C^k(E)$ is the subset of the vector space $C(E)$ of all continuous functions on E obtained by restricting the function to be k times continuously differentiable. Because sums and scalar multiples of k times continuously differentiable functions are k times continuously differentiable, then $C^k(E)$ is a subspace of $C(E)$ and a vector space in its own right.

Solution Space of a Differential Equation

The differential equation $y''-y=0$ has general solution $y=c_1e^x+c_2e^{-x}$, which means that the set S of all solutions of the differential equation consists of all possible linear combinations of the two functions e^x and e^{-x}. Briefly,

$$S = \textbf{span}\left(e^x, e^{-x}\right).$$

The functions e^x, e^{-x} are in $C^2(E)$ for any interval E on the x-axis. Therefore, S is a subspace of $C^2(E)$ and a vector space in its own right.

More generally, every homogeneous linear differential equation, of any order, has a solution set S which is a vector space in its own right.

Other Vector Spaces

The number of different vector spaces used as data storage systems in scientific literature is finite, but growing with new discoveries. There is really no limit to the number of different settings possible, because creative individuals are able to invent new settings.

Here is an example of how creation begets new vector spaces. Consider the problem $y'=2y+f(x)$ and the task of storing data for the plotting of an initial value problem with initial condition $y(x_0)=y_0$. The data set V suitable for plotting consists of column vectors

$$\vec{v} = \begin{pmatrix} x_0 \\ y_0 \\ f \end{pmatrix}.$$

A plot command takes such a data item, computes the solution

$$y(x) = y_0 e^{2x} + e^{2x} \int_0^x e^{-2t} f(t) dt$$

and then plots it in a window of fixed size with center at (x_0, y_0). The column vectors are not numerical vectors in $\mathcal{R}^3$, but some **hybrid** of vectors in $\mathcal{R}^2$ and the space of continuous functions $C(E)$ where E is the real line.

It is relatively easy to come up with definitions of vector addition and scalar multiplication on V. The closure law holds and the eight vector space properties can be routinely verified. Therefore, V is an abstract vector space, unlike any found in this text. We reiterate:

> An abstract vector space is a set V and two operations of $\boxed{+}$ and $\boxed{\cdot}$ such that the closure law holds and the eight algebraic vector space properties are satisfied.

The paycheck for having recognized a vector space setting in an application is clarity of exposition and economy of effort in details. Algebraic details in $\mathcal{R}^2$ can often be transferred unchanged to an abstract vector space setting, line for line, to obtain the details in the more abstract setting.

Independence and Dependence

The subject of *independence* applies to coordinate spaces $\mathcal{R}^n$, function spaces and in particular solution spaces of differential equations, digital photos, sequences of Fourier coefficients or Taylor coefficients, and general abstract vector spaces. Introduced here are definitions for low dimensions, the geometrical meaning of independence and basic algebraic tests for independence.

The motivation for the study of independence is the theory of general solutions, which are expressions representing *all possible solutions* of a linear problem. The subject of independence discovers *the shortest possible expression* for a general solution.

Definition 7 (Independence)
Vectors $\vec{v}_1, \ldots, \vec{v}_k$ are called **independent** provided each linear combination $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is represented by a **unique** set of constants $c_1, \ldots, c_k$.

Definition 8 (Dependence)
Vectors $\vec{v}_1, \ldots, \vec{v}_k$ are called **dependent** provided they are not independent. This means that a linear combination $\vec{v} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$ can be represented in a second way as $\vec{v} = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$ where for at least one index j, $a_j \neq b_j$.