5.4 Independence, Span and Basis

The technical topics of independence, dependence and span apply to the study of Euclidean spaces $\mathcal{R}^2$, $\mathcal{R}^3$, ..., $\mathcal{R}^n$ 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 of $A\mathbf{x} = \mathbf{0}$ by computing the partial derivatives ∂_{t_1}, ∂_{t_2}, ... of $\mathbf{x}$, where t_1, t_2, ... is the list of invented symbols assigned to the free variables, which were identified in $\mathbf{rref}(A)$. The partial derivatives are special solutions to $A\mathbf{x} = \mathbf{0}$. Knowing these special solutions is sufficient for writing out the general solution. In this sense, 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 ∂_{t_1}, ∂_{t_2}, ... The most important properties are **span** and **independence**.

Independence and Span

A list of vectors $\mathbf{v}_1$, ..., $\mathbf{v}_k$ is said to **span** a vector space V provided V is exactly the set of all linear combinations

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

The notion originates with the general solution $\mathbf{v}$ of a matrix system $A\mathbf{v} = \mathbf{0}$, where the invented symbols t_1, t_2, ... are the constants c_1, ..., c_k and the vector partial derivatives $\partial_{t_1}\mathbf{v}$, ..., $\partial_{t_k}\mathbf{v}$ are the symbols $\mathbf{v}_1$, ..., $\mathbf{v}_k$.

Vectors $\mathbf{v}_1$, ..., $\mathbf{v}_k$ are said to be **independent** provided each linear combination $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ is represented by a unique set of constants c_1, ..., c_k. See pages 301 and 304 for independence tests.

A **basis** of a vector space V is defined to be an independent set $\mathbf{v}_1$, ..., $\mathbf{v}_k$ that additionally spans V.

The 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 happens to be endowed 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 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

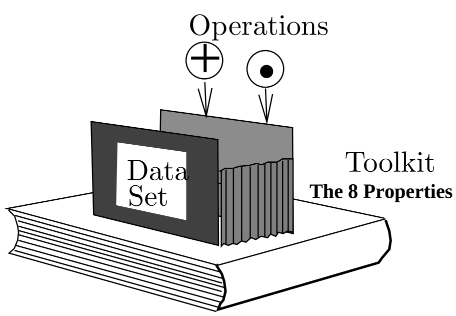

Figure 9. A Data Storage System.
A vector space is a data set of data item packages plus a storage system which organizes the data. A toolkit is provided consisting of operations + and $\cdot$ plus 8 algebraic vector space properties.

Fixed vectors and the toolkit. Scalar multiplication is a toolkit item for fixed vectors because of unit systems, like the fps, cgs and mks systems. We might originally record a data set in a fixed vector in units of meters and later find out that it should be in centimeters; multiplying the elements of a vector by the conversion factor $k = 100$ **scales** the data set to centimeters.

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

Plotting 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 $\mathbf{v}_2 = \mathbf{v}_1 + \mathbf{E}(\mathbf{v}_1)$. Symbol $\mathbf{v}_1$ is the previous dot and symbol $\mathbf{E}(\mathbf{v}_1)$ is the Euler increment. We define

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

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

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** of different colors arranged in a two dimensional array. Structure can be assigned to the photo by storing the digital data in a matrix A of size $n \times m$. Each entry of A is an integer which specifies the color properties of a given pixel.

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 separation transparencies. One easy way to do this is to code each entry of A as $a_{ij} = r_{ij} + g_{ij}x + b_{ij}x^2$ where is 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 to duplicate the original photograph.

Printing machinery used to employ separation negatives and multiple printing runs to make book photos. The advent of digital printers and better, less expensive technologies has made the separation process nearly obsolete. To help the reader understand the historical events, we record the following quote from Sam Wang[5]:

> 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

[5]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

> 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 premier storage systems used for applications involving ordinary or partial differential equations are *function spaces*. 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.

Numerical researchers in differential equations might view a function as being a fixed vector. Their unique 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 trouble with this viewpoint 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 sample set in $\mathcal{R}^{200}$ and g needs its sample set in $\mathcal{R}^{400}$. The absence of a universal numerical data storage system for sampled functions is what initially drove mathematicians and scientists to consider a storage system like the set of all functions.

Novices often 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 $f(t)$ for each possible sample t, e.g.,

$$\mathbf{f} = \begin{pmatrix} \vdots \\ f(t) \\ \vdots \end{pmatrix}.$$

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 changing from finite column vectors with entries f_1, …, f_n to function notation $f(1)$, …, $f(n)$ and then finally to $f(t)$ and a continuous variable t valid for all real numbers.

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:

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

The rules define **addition** and **scalar multiplication** of functions and immediately it is clear that the closure law for a vector space holds. Routine but long justifications are required to show that V under the above rules for addition and scalar multiplication is a vector space. This means that the 8-property toolkit is available:

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 be an open bounded set, for example $E = \{x : 0 < x < 1\}$ on the real line. The set $C(E)$ is the subset 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 $C(E)$ is a subspace of V and a vector space in its own right.

The Space $C^1(E)$. The set $C^1(E)$ is the subset of the set $C(E)$ 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 set $C(E)$ 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 $y = c_1 e^x + c_2 e^{-x}$, 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}. The latter are functions in $C^2(E)$ where E can be any interval 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 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 ones.

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 fixed vectors

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

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 fixed 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 complex setting.

Independence and Dependence

The subject of *independence* applies to coordinate spaces $\mathcal{R}^n$, function spaces and general abstract vector spaces. Introduced here are definitions for low dimensions, the geometrical meaning of independence, basic algebraic tests for independence, and generalizations to abstract vector spaces.

Definition 3 (Independence)
Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are called **independent** provided each linear combination $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ is represented by a **unique** set of constants $c_1, \ldots, c_k$.

Independence and Dependence for Two Vectors. Two vectors $\mathbf{v}_1$, $\mathbf{v}_2$ in $\mathcal{R}^2$ are said to be **independent** provided neither is the zero vector and one is not a scalar multiple of the other. Graphically, this means $\mathbf{v}_1$ and $\mathbf{v}_2$ form the edges of a non-degenerate parallelogram.

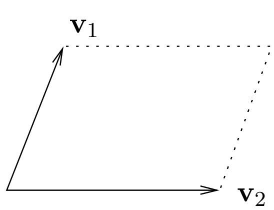

Figure 10. Independent vectors. Two nonzero nonparallel vectors $\mathbf{v}_1$, $\mathbf{v}_2$ form the edges of a parallelogram. A vector $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ lies interior to the parallelogram if the scaling constants satisfy $0 < c_1 < 1$, $0 < c_2 < 1$.

Algebraic independence test for two vectors. Given two vectors $\mathbf{v}_1$, $\mathbf{v}_2$, construct the system of equations in unknowns c_1, c_2

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}.$$

Solve the system for c_1, c_2. The two vectors are **independent** if and only if the system has the unique solution $c_1 = c_2 = 0$.

The test is equivalent to the statement that $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$ holds for one unique set of constants x_1, x_2. The details: if $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$ and also $\mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2$, then subtraction of the two equations gives $(a_1 - b_1)\mathbf{v}_1 + (a_2 - b_2)\mathbf{v}_2 = \mathbf{0}$. This is a relation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$

with $c_1 = a_1 - b_1$, $c_2 = a_2 - b_2$. Independence means $c_1 = c_2 = 0$, or equivalently, $a_1 = b_1$, $a_2 = b_2$, giving that $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$ holds for exactly one unique set of constants x_1, x_2.

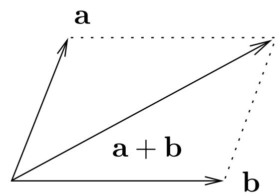

Figure 11. The parallelogram rule. Two nonzero vectors **a**, **b** are added by the parallelogram rule: $\mathbf{a}+\mathbf{b}$ has tail matching the joined tails of **a**, **b** and head at the corner of the completed parallelogram.

Why does the test work? Vector $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ is formed by the parallelogram rule, Figure 11, by adding the scaled vectors $\mathbf{a} = c_1\mathbf{v}_1$, $\mathbf{b} = c_2\mathbf{v}_2$. The zero vector $\mathbf{v} = \mathbf{0}$ can be obtained from nonzero nonparallel vectors $\mathbf{v}_1$, $\mathbf{v}_2$ only if the scaling factors c_1, c_2 are both zero.

Dependence of two vectors. Define vectors $\mathbf{v}_1$, $\mathbf{v}_2$ in $\mathcal{R}^2$ to be **dependent** provided they are **not independent**. This means one of $\mathbf{v}_1$, $\mathbf{v}_2$ is the zero vector or else $\mathbf{v}_1$ and $\mathbf{v}_2$ lie along the same line: *the two vectors cannot form a parallelogram.* Algebraic detection of dependence is by failure of the independence test: after solving the system $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$, one of the two constants c_1, c_2 is nonzero.

Independence and Dependence of Two Vectors in an Abstract Space. The algebraic definition used for $\mathcal{R}^2$ is invoked to define independence of two vectors in an abstract vector space. An immediate application is in $\mathcal{R}^3$, where all the geometry discussed above still applies. In other spaces, the geometry vanishes, but algebra remains a basic tool.

> **Independence test for two vectors $\mathbf{v}_1$, $\mathbf{v}_2$.** In an abstract vector space V, form the equation
>
> $$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}.$$
>
> Solve this equation for c_1, c_2. Then $\mathbf{v}_1$, $\mathbf{v}_2$ are independent in V only if the system has unique solution $c_1 = c_2 = 0$.

It is not obvious how to solve for c_1, c_2 in the algebraic independence test in a function space. This algebraic problem is a subject of the examples.

Illustration. Two column vectors are tested for independence by forming the system of equations $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$, e.g,

$$c_1\begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is a homogeneous system $A\mathbf{c} = \mathbf{0}$ with

$$A = \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The system $A\mathbf{c} = \mathbf{0}$ can be solved for $\mathbf{c}$ by **rref** methods. Because $\mathbf{rref}(A) = I$, then $c_1 = c_2 = 0$, which verifies independence.

If A is square and $\mathbf{rref}(A) = I$, then A^{-1} exists. The equation $A\mathbf{c} = \mathbf{0}$ has solution $\mathbf{c} = A^{-1}\mathbf{0} = \mathbf{0}$, which means $c_1 = c_2 = 0$. Theory says $A^{-1} = \mathbf{adj}(A)/\det(A)$ exists precisely when $\det(A) \neq 0$, therefore independence is verified independently of **rref** methods by the 2×2 determinant computation $\det(A) = -3 \neq 0$.

Remarks about $\det(A)$ apply to independence testing for any two vectors, but only in case the system of equations $A\mathbf{c} = \mathbf{0}$ is square. For instance, in $\mathcal{R}^3$, the homogeneous system

$$c_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has vector-matrix form $A\mathbf{c} = \mathbf{0}$ with 3×2 matrix A. There is **no chance to use determinants**. We remark that **rref** methods apply as before to verify independence.

Independence and Dependence for Three Vectors.

Following the ideas of the preceding paragraph, three vectors in $\mathcal{R}^3$ are said to be independent provided none of them are the zero vector and they form the edges of a non-degenerate parallelepiped of positive volume. Such vectors are called a **triad**. In the special case of all pairs orthogonal (the vectors are 90° apart) they are called an **orthogonal triad**.

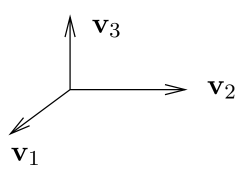

Figure 12. Independence of three vectors. Vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ form the edges of a parallelepiped. Vectors $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ satisfying $0 < c_i < 1$, $i = 1, 2, 3$, are located interior to the parallelepiped.

Independence test for three vectors. Given three vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, construct the system of equations in unknowns c_1, c_2, c_3

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}.$$

Solve the system for c_1, c_2, c_3. The vectors are **independent** if and only if the system has unique solution $c_1 = c_2 = c_3 = 0$.

Why does the test work? The vector $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ is formed by two applications of the parallelogram rule: first add the scaled vectors $c_1\mathbf{v}_1$, $c_2\mathbf{v}_2$ and secondly add the scaled vector $c_3\mathbf{v}_3$ to the resultant. The zero vector $\mathbf{v} = \mathbf{0}$ can be obtained from a vector triad $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ only if the scaling factors c_1, c_2, c_3 are all zero.

Dependence of three vectors. Given vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, they are **dependent** if and only if they are **not independent**. Then one of them is the zero vector or else one of them is in the plane of the other two, or else two of them lie along the same line. In short, three dependent vectors in $\mathcal{R}^3$ cannot be the edges of a parallelepiped. Algebraic detection of dependence is by failure of the independence test: after solving the system $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$, one of the three constants c_1, c_2, c_3 is nonzero[6].

Independence in an Abstract Vector Space. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be a finite set of vectors in an abstract vector space V. The set is **independent** if and only if the system of equations in unknowns $c_1, \ldots, c_k$

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

has unique solution $c_1 = \cdots = c_k = 0$.

Independence means that each linear combination $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ is represented by a unique set of constants $c_1, \ldots, c_k$.

A set of vectors is called **dependent** if and only if it is not independent. This means that the system of equations in variables $c_1, \ldots, c_k$ has a solution with at least one variable nonzero. In particular, any finite set of vectors containing the zero vector $\mathbf{0}$ has to be dependent, and two vectors are dependent precisely when one is a scalar multiple of the other.

Independence and Dependence Tests

Recorded here are a number of useful algebraic tests to determine independence or dependence of a finite list of vectors. The first two tests are designed for fixed vectors, while the remaining ones are designed for vector spaces of functions.

Rank Test. In the vector space $\mathcal{R}^n$, the key to detection of independence is **zero free variables**, or nullity zero, or equivalently, maximal rank. The test is justified from the formula $\mathbf{nullity}(A) + \mathbf{rank}(A) = k$, where k is the column dimension of A.

Theorem 13 (Rank-Nullity Test)
Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be k column vectors in $\mathcal{R}^n$ and let A be the augmented matrix of these vectors. The vectors are independent if $\mathbf{rank}(A) = k$ and dependent if $\mathbf{rank}(A) < k$. The conditions are equivalent to $\mathbf{nullity}(A) = 0$ and $\mathbf{nullity}(A) > 0$, respectively.

[6] In practical terms, there is at least one free variable, or equivalently, appearing in the parametric solution is at least one parameter t_1, t_2, ...

Determinant Test. In the unusual case when the system arising in the independence test can be expressed as $A\mathbf{c} = \mathbf{0}$ and A is square, then $\det(A) = 0$ detects dependence, and $\det(A) \neq 0$ detects independence. The reasoning is based upon the formula $A^{-1} = \mathbf{adj}(A)/\det(A)$, valid exactly when $\det(A) \neq 0$.

Theorem 14 (Determinant Test)
Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be n column vectors in $\mathcal{R}^n$ and let A be the augmented matrix of these vectors. The vectors are independent if $\det(A) \neq 0$ and dependent if $\det(A) = 0$.

Sampling Test for Functions. Let $f_1, \ldots, f_k$ be functions on a domain D. Let V be the vector space of all functions on D with the usual scalar multiplication and addition rules learned in college algebra. Addressed here is the question of how to test independence and dependence of $f_1, \ldots, f_k$ in V. The relation

$$c_1 f_1 + c_2 f_2 + \cdots + c_k f_k = 0$$

means

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_k f_k(x) = 0, \quad x \text{ in } D.$$

An idea how to solve for $c_1, \ldots, c_k$ arises by **sampling**, which means k relations are obtained by choosing k values for x, say $x_1, \ldots, x_k$. The equations arising are

$$\begin{array}{ccccccccc} c_1 f_1(x_1) & + & c_2 f_2(x_1) & + & \cdots & + & c_k f_k(x_1) & = & 0, \\ c_1 f_1(x_2) & + & c_2 f_2(x_2) & + & \cdots & + & c_k f_k(x_2) & = & 0, \\ \vdots & & & \vdots & \cdots & \vdots & & & \vdots \\ c_1 f_1(x_k) & + & c_2 f_2(x_k) & + & \cdots & + & c_k f_k(x_k) & = & 0. \end{array}$$

The unknowns are $c_1, \ldots, c_k$ and the coefficient matrix is

$$A = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \cdots & \vdots \\ f_1(x_k) & f_2(x_k) & \cdots & f_k(x_k) \end{pmatrix}$$

The system $A\mathbf{c} = \mathbf{0}$ has unique solution $\mathbf{c} = \mathbf{0}$, proving $f_1, \ldots, f_k$ independent, provided $\det(A) \neq 0$. It is **not true** that independence of the functions implies $\det(A) \neq 0$; it depends on the values used for the sampling.

Vandermonde Determinant. Choosing the functions in the *sampling test* to be the usual polynomial basis $f_1(x) = 1$, $f_2(x) = x, \ldots, f_k(x) =$

x^{k-1} gives the **Vandermonde matrix**

$$V(x_1,\ldots,x_k) = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{k-1} \\ 1 & x_2 & \cdots & x_2^{k-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_k & \cdots & x_k^{k-1} \end{pmatrix}.$$

Let us prove that $\det(V) \neq 0$ for distinct values $x_1, \ldots, x_k$, by establishing the identity

$$\det(V(x_1,\ldots,x_k)) = \prod_{i<j} (x_j - x_i).$$

The identity is proved from determinant properties, as follows. Let $f(x) = \det(V(x, x_2, \ldots, x_k)$. Duplicate rows in a determinant cause it to have zero value, therefore the polynomial $f(x)$ has roots $x_2, \ldots, x_k$. The factor theorem of college algebra applies to give the formula

$$f(x) = c(x_2 - x)\cdots(x_k - x),$$

where c is a constant. Cofactor expansion along the first row and matching of the coefficient of $(-x)^k$ shows that $c = \det(V(x_2, \ldots, x_k)$ [we used $(-1)^{k+1+1}(-1)^k = 1$]. Then

$$\det(V(x_1, x_2, \ldots, x_k)) = \det(V(x_2, \ldots, x_k))\Pi_{j=2}^k (x_j - x_1).$$

The relation is solved recursively to give the claimed formula. For example,

$$\begin{aligned} \det(V(x_1,x_2,x_3)) &= \det(V(x_2,x_3))(x_2 - x_1)(x_3 - x_1) \\ &= [\det(V(x_3))(x_3 - x_2)](x_2 - x_1)(x_3 - x_1) \\ &= (x_3 - x_2)(x_2 - x_1)(x_3 - x_1). \end{aligned}$$

Wronskian Test for Functions. Given functions $f_1, \ldots, f_n$ each differentiable $n-1$ times on an interval $a < x < b$, the **Wronskian determinant** [7] is defined by the relation

$$W(f_1,\ldots,f_n)(x) = \det \begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \cdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{pmatrix}.$$

Theorem 15 (Wronskian Test)
Let functions $f_1, \ldots, f_n$ be differentiable $n-1$ times on interval $a < x < b$. Then $W(f_1, \ldots, f_n)(x_0) \neq 0$ for some x_0 in (a, b) implies $f_1, \ldots, f_n$ are independent functions in the vector space V of all functions on (a, b).

[7] Named after mathematician Jósef Maria Hoëné Wronski (1778-1853), born in Poland.

Proof: The objective of the proof is to solve the equation

$$c_1 f_1 + c_2 f_2 + \cdots + c_n f_n = 0$$

for the constants $c_1, \ldots, c_n$, showing they are all zero. The idea of the proof, attributed to Wronski, is to differentiate the above equation $n-1$ times, then substitute $x = x_0$ to obtain a homogeneous $n \times n$ system $A\mathbf{c} = \mathbf{0}$ for the components $c_1, \ldots, c_n$ of the vector $\mathbf{c}$. Because $\det(A) = W(f_1, \ldots, f_n)(x_0) \neq 0$, Cramer's rule applies to show that $\mathbf{c} = \mathbf{0}$, completing the proof.

Exercises 5.4

General Solution.

1.

Independence and Span.

2.

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

3.

Digital Photographs.

4.

Function Spaces.

5.

The Space $C(E)$.

6.

The Space $C^1(E)$.

7.

The Space $C^k(E)$.

8.

Solution Space.

9.

Independence and Dependence.

10.

Algebraic Independence Test for Two Vectors.

11.

Dependence of two vectors.

12.

13.

Independence Test for Three Vectors.

14.

Dependence of Three Vectors.

15.

Independence in an Abstract Vector Space.

16.

Rank Test.

17.

18.

Determinant Test.

19.

Sampling Test for Functions.

20.

Vandermonde Determinant.

21.

Wronskian Test for Functions.

22.

23.

5.5 Basis, Dimension and Rank

The topics of basis, dimension and rank apply to the study of Euclidean spaces, continuous function spaces, spaces of differentiable functions and general abstract vector spaces.

A **basis** for a vector space V is defined to be an independent set of vectors whose finite linear combinations span V. If the set of independent vectors is finite, then V is called **finite dimensional** and otherwise it is called **infinite dimensional**. The **dimension** of V is the number of vectors in a basis. Because of the following result, for finite dimensional V, the term *dimension* is well-defined.

Theorem 16 (Dimension)
If a vector space V has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_p$ and also a basis $\mathbf{u}_1, \ldots, \mathbf{u}_q$, then $p = q$.

Proof: Assume the hypotheses true and the conclusion false, in order to reach a contradiction. Let the larger basis be listed first, $p > q$. Because $\mathbf{u}_1, \ldots, \mathbf{u}_q$ is a basis of V, then there are coefficients $\{a_{ij}\}$ such that

$$\mathbf{v}_i = a_{i1}\mathbf{u}_1 + \cdots + a_{iq}\mathbf{u}_q, \quad 1 \le i \le p.$$

Let $A = [a_{ij}]$ be the $p \times q$ matrix of coefficients. Because $p > q$, then $\mathbf{rref}(A^T)$ has at most q leading variables and at least $p - q > 0$ free variables. Then the $q \times p$ homogeneous system $A^T\mathbf{x} = \mathbf{0}$ has infinitely many solutions. Let $\mathbf{x}$ be a nonzero solution of $A^T\mathbf{x} = \mathbf{0}$. The equation $A^T\mathbf{x} = \mathbf{0}$ means $\sum_{i=1}^p a_{ij}x_i = 0$ for $1 \le j \le p$, giving the dependence relation

$$\begin{array}{rcl} \sum_{i=1}^p x_i\mathbf{v}_i & = & \sum_{i=1}^p x_i \sum_{j=1}^q a_{ij}\mathbf{u}_j \\ & = & \sum_{j=1}^q \sum_{i=1}^p a_{ij}x_i\mathbf{u}_j \\ & = & \sum_{j=1}^q (0)\mathbf{u}_j \\ & = & \mathbf{0} \end{array}$$

The independence of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is contradicted, completing the proof.

Euclidean Spaces. The space $\mathcal{R}^n$ has a **standard basis** consisting of the columns of the $n \times n$ identity matrix, therefore $\mathcal{R}^n$ has dimension n. More generally,

Theorem 17 (Bases in $\mathcal{R}^n$)
Any basis of $\mathcal{R}^n$ has exactly n independent vectors. Further, any list of $n + 1$ or more vectors in $\mathcal{R}^n$ is dependent.

Polynomial Spaces. The vector space of all polynomials $p(x) = p_0 + p_1x + p_2x^2$ has dimension 3, because a basis is 1, x, x^2 in this function space. Formally, the basis elements are obtained from the general solution $p(x)$ by partial differentiation on the symbols p_0, p_1, p_2.

Differential Equations. The equation $y'' + y = 0$ has general solution $y = c_1 \cos x + c_2 \sin x$. Therefore, the formal partial derivatives ∂_{c_1}, ∂_{c_2}

applied to the general solution y give a basis $\cos x$, $\sin x$. The solution space of $y'' + y = 0$ has dimension 2.

Similarly, $y''' = 0$ has a solution basis 1, x, x^2 and therefore its solution space has dimension 3. Generally, an nth order linear homogeneous scalar differential equation has solution space V of dimension n, and an $n \times n$ linear homogeneous system $\mathbf{y}' = A\mathbf{y}$ has solution space V of dimension n. A general procedure for finding a basis for a differential equation:

> Let a differential equation have general solution expressed in terms of arbitrary constants c_1, c_2, ..., then a basis is found by taking the partials ∂_{c_1}, ∂_{c_2}, ...

Largest Subset of Independent Vectors

Let vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be given in $\mathcal{R}^n$. The subset V of $\mathcal{R}^n$ consisting of all linear combinations $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ is closed under addition and scalar multiplication. In short, the set V is a subspace of $\mathcal{R}^n$.

Discussed here are efficient methods for finding a basis for V. Equivalently, we find a *largest subset* of independent vectors from the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Such a largest subset spans V and is independent, therefore it is a basis for V.

An Iterative Method. A largest independent subset can be identified as $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_p}$ for some distinct subscripts $i_1, \ldots, i_p$. We describe how to find such subscripts. Let i_1 be the first subscript such that $\mathbf{v}_{i_1} \neq \mathbf{0}$. Define i_2 to be the *first* subscript greater than i_1 such that

$$\mathbf{rank}(\mathbf{aug}(\mathbf{v}_1, \ldots, \mathbf{v}_{i_1})) < \mathbf{rank}(\mathbf{aug}(\mathbf{v}_1, \ldots, \mathbf{v}_{i_2})).$$

The process terminates if there is no such $i_2 > i_1$. Otherwise, proceed in a similar way to define i_3, i_4, ..., i_p. The rank test uses the basic tools of **swap**, **combination** and **multiply**. An efficient shortcut is the following:

> The rank of $\mathbf{aug}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i_q})$ is the same as the rank of $\mathbf{aug}(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_q})$.

The shortcut allows the **rref** tools to be applied to smaller matrices, to the same effect.

Why does it work? Because each column added to the augmented matrix which increases the rank cannot be a linear combination of the preceding columns. In short, that column is independent of the preceding columns.

Pivot Column Method. A column j of A is called a **pivot column** provided $\mathbf{rref}(A)$ has a leading one in column j. The leading ones in $\mathbf{rref}(A)$ belong to consecutive columns of the identity matrix I.

Lemma 1 (Pivot Columns and Dependence) A non-pivot column of A is a linear combination of the pivot columns of A.

Proof: Let column j of A be non-pivot. Consider the homogeneous system $A\mathbf{x} = \mathbf{0}$. The pivot columns subscripts determine the leading variables and the remaining column subscripts determine the free variables. Define $x_j = 1$. Define all other free variables to be zero. The lead variables are now determined and the resulting nonzero vector $\mathbf{x}$ satisfies the homogeneous equation $A\mathbf{x} = \mathbf{0}$. Translating this equation into a linear combination of columns says that column j is a linear combination of the pivot columns of A.

Theorem 18 (Independence)
The pivot columns of a matrix A are linearly independent.

Proof: Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be the columns of A and let $i_1, \ldots, i_p$ be the pivot columns of A. Independence is proved by solving the system of equations

$$c_1\mathbf{v}_{i_1} + \cdots + c_p\mathbf{v}_{i_p} = \mathbf{0}$$

for the constants $c_1, \ldots, c_p$, eventually verifying they are all zero. The tool used to solve for the constants is the formula

$$A = E_1E_2 \cdots E_r\,\mathbf{rref}(A),$$

where $E_1, \ldots, E_r$ denote certain elementary matrices. Each matrix is the inverse of a swap, multiply or combination operation applied to A, in order to reduce it to **rref** form. Because $A = \mathbf{aug}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, then

$$\mathbf{v}_{i_q} = E_1E_2 \cdots E_r\mathbf{e}_{i_q}, \quad q = 1, \ldots, p,$$

where $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_p}$ denote distinct columns of the identity matrix, which occupy the columns of the leading ones in $\mathbf{rref}(A)$. Then

$$\begin{aligned} \mathbf{0} &= c_1\mathbf{v}_{i_1} + \cdots + c_p\mathbf{v}_{i_p} \\ &= E_1E_2 \cdots E_r(c_1\mathbf{e}_{i_1} + \cdots + c_p\mathbf{e}_{i_p}) \end{aligned}$$

implies by invertibility of elementary matrices that

$$c_1\mathbf{e}_{i_1} + \cdots + c_p\mathbf{e}_{i_p} = \mathbf{0}.$$

However, distinct columns of the identity matrix are independent, therefore $c_1 = \cdots = c_p = 0$. The independence of the pivot columns of A is established.

Rank and Nullity

The **rank** of a matrix A equals the number of leading ones in $\mathbf{rref}(A)$. The **nullity** of a matrix A is the number of columns of A less the rank of A. Symbols $\mathbf{rank}(A)$ and $\mathbf{nullity}(A)$ denote these two integer values and

$$\mathbf{rank}(A) + \mathbf{nullity}(A) = \text{column dimension of } A.$$

In terms of the system $A\mathbf{x} = \mathbf{0}$, the rank of A is the number of leading variables and the nullity of A is the number of free variables, in the reduced echelon system $\mathbf{rref}(A)\mathbf{x} = \mathbf{0}$.

Theorem 19 (Basis for $A\mathbf{x} = \mathbf{0}$)
Assume $k = \mathbf{nullity}(A) > 0$. Then the solution set of $A\mathbf{x} = \mathbf{0}$ can be expressed as

$$\mathbf{x} = t_1\mathbf{x}_1 + \cdots + t_k\mathbf{x}_k \tag{1}$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent solutions of $A\mathbf{x} = \mathbf{0}$ and $t_1, \ldots, t_k$ are arbitrary scalars. The meaning:

$$\mathbf{nullity}(A) = \dim\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}.$$

Proof: The system $\mathbf{rref}(A)\mathbf{x} = \mathbf{0}$ has exactly the same solution set as $A\mathbf{x} = \mathbf{0}$. This system has a standard general solution $\mathbf{x}$ expressed in terms of invented symbols $t_1, \ldots, t_k$. Define $\mathbf{x}_j = \partial_{t_j}\mathbf{x}$, $j = 1, \ldots, k$. Then (1) holds. It remains to prove independence, which means we are to solve for $c_1, \ldots, c_k$ in the system

$$c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = \mathbf{0}.$$

The left side is a solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{0}$ in which the invented symbols have been assigned values $c_1, \ldots, c_k$. The right side implies each component of $\mathbf{x}$ is zero. Because the standard general solution assigns invented symbols to free variables, the relation above implies that each free variable is zero. But free variables have already been assigned values $c_1, \ldots, c_k$. Therefore, $c_1 = \cdots = c_k = 0$. The proof is complete.

Theorem 20 (Row Rank Equals Column Rank)
The number of independent rows of a matrix A equals the number of independent columns of A. Equivalently, $\mathbf{rank}(A) = \mathbf{rank}(A^T)$.

Proof: Let C be the set of all linear combinations of columns of A. The non-pivot columns of A are linear combinations of pivot columns of A. Therefore, any linear combination of columns of A is a linear combination of the $p = \mathbf{rank}(A)$ linearly independent pivot columns. By definition, the pivot columns form a **basis** for the vector space C.

Let R be the set of all linear combinations of columns of A^T (the rows of A). Let q be the number of elements in a basis for R. It will be shown that $p = q$, which proves the theorem.

Let $\mathbf{rref}(A) = E_1 \cdots E_k A$ where $E_1, \ldots, E_k$ are elementary swap, multiply and combination matrices. Then $E = (E_1 \cdots E_k)^T$ is invertible and $\mathbf{rref}(A)^T = A^T E$, $\mathbf{rref}(A)^T E^{-1} = A^T$. The matrix $\mathbf{rref}(A)^T$ has its first p columns independent and its remaining columns are zero. Each nonzero column of $\mathbf{rref}(A)^T$ is expressed on the right as a linear combination of the columns of A^T. Therefore, R contains p independent vectors. This gives $p \le q$.

Because $\mathbf{rref}(A)^T E^{-1} = A^T$, every column of A^T is a linear combination of the p nonzero columns of $\mathbf{rref}(A)^T$. This implies a basis for R contains at most p elements, i.e., $q \le p$, which proves $p = q$, and completes the proof of the theorem.

The results of the preceding theorems are combined to obtain the **pivot method** for finding a largest independent subset.

Theorem 21 (Pivot Method)
Let A be the augmented matrix of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Let the leading ones in $\mathbf{rref}(A)$ occur in columns $i_1, \ldots, i_p$. Then a largest independent subset of the k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is the set

$$\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_p}.$$

Nullspace, Column Space and Row Space

The **kernel** or **nullspace** of an $m \times n$ matrix A is the vector space of all solutions $\mathbf{x}$ to the homogeneous system $A\mathbf{x} = \mathbf{0}$. In symbols,

$$\mathbf{kernel}(A) = \mathbf{nullspace}(A) = \{\mathbf{x} \ : \ A\mathbf{x} = \mathbf{0}\}.$$

The **column space** of $m \times n$ matrix A is the vector space consisting of all vectors $\mathbf{y} = A\mathbf{x}$, where $\mathbf{x}$ is arbitrary in $\mathcal{R}^n$. Algebra texts might also call the column space the **image** of A. Because $A\mathbf{x}$ can be written as a linear combination of the columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of A, the column space is the set of all linear combinations

$$\mathbf{y} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n.$$

In symbols,

$$\mathbf{Image}(A) = \mathbf{colspace}(A) = \{\mathbf{y} \ : \ \mathbf{y} = A\mathbf{x} \text{ for some } \mathbf{x}\}.$$

The **row space** of $m \times n$ matrix A is the vector space consisting of vectors $\mathbf{w} = A^T\mathbf{y}$, where $\mathbf{y}$ is arbitrary in $\mathcal{R}^m$. Technically, the row space of A is the column space of A^T. This vector space is viewed as the set of all linear combinations of rows of A. In symbols,

$$\mathbf{rowspace}(A) = \mathbf{colspace}(A^T) = \{\mathbf{w} \ : \ \mathbf{w} = A^T\mathbf{y} \text{ for some } \mathbf{y}\}.$$

The row space of A and the null space of A live in $\mathcal{R}^n$, but the column space of A lives in $\mathcal{R}^m$. The correct bases are obtained as follows. If an alternative basis for $\mathbf{rowspace}(A)$ is suitable (rows of A not reported), then bases for $\mathbf{rowspace}(A)$, $\mathbf{colspace}(A)$, $\mathbf{nullspace}(A)$ can all be found by calculating just $\mathbf{rref}(A)$.

Null Space. Compute $\mathbf{rref}(A)$. Write out the general solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{0}$, where the free variables are assigned parameter names t_1, ..., t_k. Report the basis for $\mathbf{nullspace}(A)$ as the list of partials $\partial_{t_1}\mathbf{x}$, ..., $\partial_{t_k}\mathbf{x}$.

Column Space. Compute $\mathbf{rref}(A)$. Identify the lead variable columns i_1, ..., i_k. Report the basis for $\mathbf{colspace}(A)$ as the list of columns i_1, ..., i_k of A.

Row Space. Compute $\mathbf{rref}(A^T)$. Identify the lead variable columns i_1, ..., i_k. Report the basis for $\mathbf{rowspace}(A)$ as the list of rows i_1, ..., i_k of A.

Alternatively, compute $\mathbf{rref}(A)$, then $\mathbf{rowspace}(A)$ has a basis consisting of the list of nonzero rows of $\mathbf{rref}(A)$. The two bases obtained by these methods are different, but equivalent.

Due to the identity $\mathbf{nullity}(A) + \mathbf{rank}(A) = n$, where n is the column dimension of A, the following results hold. Notation: $\dim(V)$ is the dimension of vector space V, which equals the number of elements in a basis for V. Recall that $\mathbf{nullspace}(A) = \mathbf{kernel}(A)$ and $\mathbf{colspace}(A) = \mathbf{Image}(A)$ are subspaces with dual naming conventions in the literature.

Theorem 22 (Dimension Identities)

(a) $\dim(\mathbf{nullspace}(A)) = \dim(\mathbf{kernel}(A)) = \mathbf{nullity}(A)$

(b) $\dim(\mathbf{colspace}(A)) = \dim(\mathbf{Image}(A)) = \mathbf{rank}(A)$

(c) $\dim(\mathbf{rowspace}(A)) = \mathbf{rank}(A)$

(d) $\dim(\mathbf{kernel}(A)) + \dim(\mathbf{Image}(A)) = \text{column dimension of } A$

(e) $\dim(\mathbf{kernel}(A)) + \dim(\mathbf{kernel}(A^T)) = \text{column dimension of } A$

Equivalent Bases

Assume $\mathbf{v}_1$, ..., $\mathbf{v}_k$ are independent vectors in an abstract vector space V and S is the subspace of V consisting of all linear combinations of $\mathbf{v}_1$, ..., $\mathbf{v}_k$.

Let $\mathbf{u}_1$, ..., $\mathbf{u}_\ell$ be independent vectors in V. We study the question of whether or not $\mathbf{u}_1$, ..., $\mathbf{u}_\ell$ is a basis for S. First of all, all the vectors $\mathbf{u}_1$, ..., $\mathbf{u}_\ell$ have to be in S, otherwise this set cannot possibly span S. Secondly, to be a basis, the vectors $\mathbf{u}_1$, ..., $\mathbf{u}_\ell$ must be independent. Two bases for S must have the same number of elements, by Theorem 16. Therefore, $k = \ell$ is necessary for a possible second basis of S.

Theorem 23 (Equivalent Bases of a Subspace S)
Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be independent vectors in an abstract vector space V. Let S be the subspace of V consisting of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. A set of vectors $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ in V is an equivalent basis for S if and only

(1) Each of $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

(2) The set $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ is independent.

(3) The sets are the same size, $k = \ell$.

Practical tests for column vectors in $\mathcal{R}^n$ can be constructed for computer algebra systems and numerical laboratories. Define

$$\begin{aligned} B &= \mathbf{aug}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \\ C &= \mathbf{aug}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell) \\ W &= \mathbf{aug}(B, C) \end{aligned}$$

Then test the relation

$$k = \ell = \mathbf{rank}(B) = \mathbf{rank}(C) = \mathbf{rank}(W).$$

The relation implies that the two sets of vectors are independent and span the same space. The following `maple` code implements these ideas to verify that the two bases determined from the `colspace` command in `maple` and the pivot columns of A are equivalent. In `maple`, the report of the column space basis is identical to the nonzero rows of $\mathbf{rref}(A^T)$.

```
with(linalg):
A:=matrix([[1,0,3],[3,0,1],[4,0,0]]);
colspace(A);          # Solve Ax=0, basis v1,v2 below
v1:=vector([2,0,-1]);v2:=vector([0,2,3]);
rref(A);              # Find the pivot cols=1,3
u1:=col(A,1); u2:=col(A,3); # pivot col basis
B:=augment(v1,v2); C:=augment(u1,u2);
W:=augment(B,C);
rank(B),rank(C),rank(W); # all equal 2
```

Exercises 5.5

Basis and Dimension.

1.

Euclidean Spaces.

2.

3.

Polynomial Spaces.

4.

Differential Equations.

5.

6.

Largest Subset of Independent Vectors.

7.

8.

9.

Row and Column Rank.

10.

Pivot Method.

11.

Nullspace, Column Space and Row Space.

12.

13.

14.

15.

Dimension Identities.

16.

Equivalent Bases.

17.

18.

19.