Matrices

OFFENBARKEIT

April 3, 2024

10 Matrices

Along with **symmetry groups**, **matrix groups** are standard objects in Galois Representation Theory. Matrix multiplication is the group operator. An $\mathcal{R}$-matrix consists only of elements from some number system $\mathcal{R}$. Examples are a $\mathbb{Z}$-matrix, $\overline{\mathbb{Q}}$-matrix, ... Matrices are rectangular, consisting of c columns and r rows, normally written enclosed in brackets.

10.1 Motivation

Linear algebra is the study of equations of lines (polynomial equations of degree 1). Matrices are used to represent systems of linear equations.

10.2 Matrix Multiplication

To multiply 2 matrices, $M = M_1 M_2$, the columns of the M_1 must number the same as the rows of M_2. An $r_1 \times n$ matrix times an $n \times c_2$ matrix results in an $r_1 \times c_2$ matrix. Multiplication proceeds by forming the dot product of rows of M_1 by columns of M_2. The dot product of two n-tuples is the sum of the pairwise products of elements. The $M_{r_i c_j}$ entry of the product matrix is the dot product of the r_ith row of M_1 and the c_jth column of M_2.

The following examples are $\mathbb{Z}$-matrices under mutiplication.

$$\begin{bmatrix} 1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 9 \\ 1 \end{bmatrix} = 1 \cdot 2 + 3 \cdot 9 + 5 \cdot 1 = 34$$

$$\begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 52 \\ 32 \end{bmatrix}$$

10.3 Solving Systems of Equations

Consider the two equations:

$3x - 5y = 2$

$2x + 3y = 14$

Expressed as matrices,

$AZ = B, \text{where A} = \begin{bmatrix} 3 & -5 \\ 2 & 3 \end{bmatrix}, \text{B} = \begin{bmatrix} 2 \\ 14 \end{bmatrix}, \text{Z} = \begin{bmatrix} x \\ y \end{bmatrix}$

$AZ = B$ has a solution if A is invertible, meaning A^{-1} exists and $AA^{-1} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Without showing the computation of A^{-1}, the solution to the two equations is found by:

$Z = A^{-1}B = \begin{bmatrix} \frac{3}{19} & \frac{5}{19} \\ \frac{-2}{19} & \frac{3}{19} \end{bmatrix} \begin{bmatrix} 2 \\ 14 \end{bmatrix} = \begin{bmatrix} \frac{76}{19} \\ \frac{38}{19} \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ Thus x=4, y=2

is a solution to the equations. The denominator 19 in the entries for the inverse matrix is a special value assigned to A, called its determinant, a number in $\mathcal{R}$. If the determinant of a matrix has an inverse in $\mathcal{R}$, then the matrix is invertible. The determinant of a matrix is used to calculate the entries in the inverse matrix when it exists. Invertibility is discussed in the next section.

10.4 Slight Digression: Basic Notions of Vector Spaces

The rows and columns of a matrix are vectors, in the sense of n-tuples with a defined arithmetic. Going beyond numbers and individual matrix computations, the theory of solutions of linear equations is based in collections of vectors, called vector spaces. The machinery associated with vector spaces allows the underlying concepts to be unified and visualized. Most of the machinery needed is also found in a generalization of a vector space called a module, which is a topic linked to general ring theory. But for the purpose here, discussion is in terms of vector spaces, which require only the machinery of fields already introduced.

Staying with the notion of vectors as n-tuples, given a field F, the set of all n-tuples of elements of F, $n \geq 1$, form a vector space F^n, aka $V^n(F)$, of dimension* $n$over F, and each n-tuple element of F^n is called a vector. (*Note this is not a definition of the dimension of a vector space, but rather an equivalent usage for coordinate spaces such as F^n.)

The vectors in F^n can be added, where addition is associative and commutative, where additive inverse vectors exist for each vector, and there is an additive identity vector. Vectors can be multiplied by numbers from F, also called scalars, where multiplication is associative, scalar multiplication distrubutes over vector addition, vectors distribute over scalar addition, and the number 1 in F leaves a vector unchanged under multiplication.

Given a vector space V, a subset of vectors in V can also form a vector space, called a subspace of V. Typically, a subspace arises as the image or kernel of a linear mapping on a vector space, or as a direct sum or quotient of two subspaces (not further defined here). The orthogonal complement of a vector space V is the space of all vectors perpendicular to all the vectors in V, $V^{\perp}$.

Examples of vector spaces follow. Any sort of objects, for which the rules of vector arithmetic apply, can become a vector space.

The trivial vector space is simply the zero vector, $\mathbf{0}$. $\mathbb{R}$ is itself a vector space of dimension 1 over $\mathbb{R}$.

$\mathbb{C}$ is a vector space of dimension 2 over $\mathbb{R}$, and more generally, a field extension forms a vector space over the extended field, using the arithmetic of that field. E.g. $\mathbb{R}$ is a vector space of uncountable dimension over $\mathbb{Q}$. $\mathbb{R}^n$ is a vector space of dimension n over $\mathbb{R}$. In particular, when $n = 3$, the vectors correspond to the set of all points in Euclidean space. In $\mathbb{R}^3$, subspaces are $\mathbf{0}$, the lines through the origin, the planes through the origin, and $\mathbb{R}^3$ itself. The set of all $m \times n$ matrices over F form a vector space of dimension mn over F. The set of all functions, continuous and single-valued on the closed unit interval, is an uncountable infinite-dimensional vector space. Polynomials over F of degree n or less form an n-dimensional vector space over F. If polynomials of all degrees are considered, the dimension is countably infinite. Pertinent to the current discussion, solutions of a set of homogeneous linear equations form a vector space over $\mathbb{F}$, because adding such solutions together or multiplying by elements of F yields further solutions.

Given vectors, $\mathbf{v}_i \in F^n$, their linear combination is expressed by

$a_1\mathbf{v}_1 + a_2 v_2 + \ldots + a_n\mathbf{v}_n$

where a_i are scalars from F. All linear combinations of vectors in some subset of a vector space generate a subspace, and the vectors in the set are said to span the subspace.

Consider a set of vectors $\mathbf{v}_i$ which have some linear combination equal to $\mathbf{0}$, the zero vector in F^n; then the vectors in the set are said to be linearly dependent. If no such linear combination exists, the set of vectors are said to be linearly independent. Equivalently, for the matrix form of a system of equations $AZ = B$ above, the columns of A are linearly independent only if Z is the zero matrix whenever B is the zero matrix.

The transpose of a matrix A, written A^T, is the matrix $A$with its rows and columns exchanged; if the elements of A are A_{ij}, the elements of A^T are A_{ji}. The principal diagonal of an $n \times n$ matrix A goes from A_{11} to A_{nn}. The transpose of a square matrix flips the elements about the principal diagonal.

All the remaining examples come from $\mathbb{R}^3$, providing the most intuitive feel for the concepts. An $\mathbb{R}^3$ discussion generalizes to any finite-dimensional vector space and in particular F^n. $\mathbb{R}^3$ is a 3-dimensional vector space over $\mathbb{R}$, the set of all triples of real numbers (x, y, z). Consider the vectors:

$\hat{\mathbf{e}}_1 = (1, 0, 0)^T,\ \hat{\mathbf{e}}_2 = (0, 1, 0)^T,\ \hat{\mathbf{e}}_3 = (0, 0, 1)^T$

The set of all linear combinations of $\hat{\mathbf{e}}_i$ generates the entire vector space $\mathbb{R}^3$. Furthermore, the $\hat{\mathbf{e}}_i$ are linearly independent. A linearly independent set of vectors that spans a vector space is called a basis. Thus, $\hat{\mathbf{e}}_i$ are basis vectors for $\mathbb{R}^3$, or simply a basis for $\mathbb{R}^3$. Geometrically, the $\hat{\mathbf{e}}_i$ are just the $\mathbf{x}, \mathbf{y}, \mathbf{z}$ coordinate unit vectors of $\mathbb{R}^3$. The condition for orthogonality of two vectors is that their dot product is 0. The $\hat{\mathbf{e}}_i$ are thus mutually orthogonal. And because they are each of unit length, they are called an orthonormal basis for $\mathbb{R}^3$, also known as the standard basis for $\mathbb{R}^3$. Any set of linearly independent vectors in $\mathbb{R}^3$ form a basis for $\mathbb{R}^3$, and all such bases contain the same number of vectors, called the dimension of the vector space over $\mathbb{R}$, written $dim_{\mathbb{R}}$ or just dim. (For vector space F^n, the number of elements in each vector and the number of vectors in each basis are the same.) Thus the 3-dimensional vector space $\mathbb{R}^3$ has three linearly independent vectors in its standard basis. Given a vector $\mathbf{v} \in \mathbb{R}^3$, there is a unique way of writing $\mathbf{v}$ as a linear combination

of basis vectors. For example,

$\mathbf{v} = a_1\hat{\mathbf{e}}_1 + a_2\hat{\mathbf{e}}_2 + a_n\hat{\mathbf{e}}_3$

for some coefficients $a_i \in \mathbb{R}$, where the a_i are called the coordinates of $\mathbf{v}$ relative to the selected basis vectors $\hat{\mathbf{e}}_i$. For the orthonormal basis $\hat{\mathbf{e}}_i$, the coordinates are the usual coordinates of Euclidean space.

The vector space concepts above can be applied to systems of linear equations. Consider vector space F^n over a field F, together with a general system of m linear equations with n unknowns written as:

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & +\cdots+ & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & +\cdots+ & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & +\cdots+ & a_{mn}x_n & = & b_m \end{array}$$

where values of the coefficients and unknowns are in F. As shown in the prior section, there is a matrix form of these equations, written $AX = B$, but now with A of size $m \times n$.

The equivalent vector form of these equations is in form $A\mathbf{x} = \mathbf{b}$:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Initially, consider just the column space generated by the left hand vectors. The collection of all possible linear combinations of the vectors on the left-hand side span a subspace called the column space of $A, col(A) \subseteq F^m$. If the column vectors are linearly independent, $col(A) = F^m$. The equations have a solution just when $\mathbf{b}$ is in $col(A)$. If every vector in $col(A)$ has exactly one expression as a linear combination of the given left-hand vectors, then any solution is unique. In any event, $col(A)$ has a basis of linearly independent vectors that do guarantee exactly one expression; and the number of vectors in that basis (its dimension) cannot be larger than m or n, but it can be smaller. The existence of m independent vectors guarantees a solution regardless of the right-hand side, which otherwise is not guaranteed.

The null space of matrix A contains all solutions to $A\mathbf{x} = \mathbf{0}$. Referring to the $m \times n$ matrix A above, the subspaces of F^n are:

row space of A, $row(A)$, aka $col(A^T)$ null space of $A, null(A)$

and the subspaces of F^m are:

column space of A above, $col(A)$ null space of $A^T, null(A^T)$.

Any linear transformation from $F^n \to F^m$ can be represented by an $m \times n$ matrix. The above shows that A is associated with such a linear projection, specifically a mapping from the row space to the column space, and from the null space to $\mathbf{0}$, where solutions in the null space are in the kernel of map.

The following additional relations hold between the subspaces:
$null(A) = row(A)^\perp$
$null(A^T) = col(A)^\perp$
dim $col(A)$ = dim $row(A) = r$
dim $null(A)$ = dim $null(A^T) = n - r$
where dim $row(A)$ is called the rank of A.

To understand the connections between the four spaces associated with solutions of $A\mathbf{x} = \mathbf{b}$, there are different conditions to consider, corresponding to constraints on the rank of matrix A. There will be seen a close connection between solutions of the non-homogeneous equation above, and homogeneous equation $A\mathbf{x} = \mathbf{0}$. In each case where a particular solution $\mathbf{x}_p$ exists to $A\mathbf{x} = \mathbf{b}$, the total solution set can be written as:

$\{\mathbf{x}_p + \mathbf{x}_n : \mathbf{x}_n$ is any solution to A$\mathbf{x} = \mathbf{0}\}$

The non-homogeneous equation is a translation of the homogeneous equation by the vector $\mathbf{x}_p$.

The 'unique solution' case corresponds to $r = m = n$, where the columns of A are linearly independent (in other words, A is invertible and the column vectors form a basis for F^n). Then $col(A) = row(A) = F^n$ and $null(A) = null(A^T) = \mathbf{0}$. There is a unique solution: $\mathbf{x}_p + \mathbf{0}$ to the system of equations, where $\mathbf{x}_p \in row(A)$ and the only solution in $null(A)$ is $\mathbf{0}$. In Euclidean space, $\mathbf{x}_p$ is a point.

In the case $r = m < n$, A has full row rank, there at least as many unknowns as equations, A has a right inverse, and there are infinitely many solutions of the form $\mathbf{x}_p + \mathbf{x}_n$. For example, in Euclidean space with $n = 3$, if $m = 2$, the two equations describe planes and then the solution is the set of points on the line at the intersection of the two planes (because rank is 2, the planes cannot be parallel). If $m = 1$, the solution is the plane described by the equation.

In the case $r = n < m$, A has full column rank, there are fewer unknowns than equations, A has a left inverse, and there may be one solution, or none. There will be a solution $\mathbf{x}_p + \mathbf{0}$ if $\mathbf{b} \in col(A)$. Else, $\mathbf{x}_p$ does not exist (the equations are not consistent).

The final 'everything else' case has $r < m$ and $r < n$. There can be either infinite solutions, or none. Infinite solutions generally obtain if $m < n$, when $\mathbf{x}_p \in row(A)$. No solutions result if $m > n$ and $r < n$.