Linear Equations

Systems of Linear Equations

Systems of equations and row operations.

For example, consider a network of flows; nodes with a given inflow and outflow. To compute flows in this network we can use a system linear equations.

Data fitting using a polynomial. Sometimes, we want to find a function of a certain form which fits to a set of data points. To find the relevant coefficients, we can use a system of linear equations.

In general, we will take the variables in a linear equation to be $x_1, x_2, \ldots, x_n$ and the coefficients to be $a_1, \ldots, a_n$.

A finite collection of linear equations of a given set of variables is called a *system of linear equations* or a *linear system*.

$$x_1 + 5x_2 + 6x_3 = 100$$

$$x_2 - x_3 = -1$$

$$-x_1 + x_3 = 11$$

Here, despite missing x_1 and x_2 respectively, the second and third equations are still part of the same system as they implicitly have a term with a 0 coefficient.

The organisation of the above system, with all variables on the right and constants at left is the standard form of presenting a system.

A *homogenous* linear system is one where all of the constants at right are 0. These systems are easier to solve, and by solving a homogenous version of a non-homogenous system, we can find a solution to the non-homogenous variant.

A solution to a linear system is a set of values for variables that cause all equations in the system to be true.

Solving by Elimination

$$(1): \ 2x - y = 3$$

$$(2): \ x + y = 0$$

$$(2) \Rightarrow y = -x$$
$$(2)\&(1) \Rightarrow 2x - (-x) = 3 \Rightarrow 3x = 3$$
$$x = 1 \Rightarrow y = -1$$

A key to the applicability of this method is that we can divide by the coefficients, which will not always be a valid assumption. This method can be implemented algorithmically and will always either yield a solution or tell you there is none.

Matrices

Really, the variables in a linear system aren't really important; it is simply the coefficients which define their relations. A matrix, a rectangular array of numbers, can be used to store these values. A $p \times q$ matrix has p rows and q columns.

A *augmented matrix* for a linear system is the matrix formed from the coefficients in the equations and the constant terms, separated by a vertical line. For example

$$\begin{aligned} 2x - y = 3 &\Rightarrow 2x + -1y = 3 \\ x + y = 0 &\Rightarrow 1x + 1y = 0 \end{aligned} = \left[\begin{array}{cc|c} 2 & -1 & 3 \\ 1 & 1 & 0 \end{array}\right]$$

With coefficients at left and constants at right. The number of rows should be equal to the number of equations. Each column corresponds to a given variable in the equations.

We can perform some *elementary row operations* to such a matrix without changing its solutions. These are

- Interchanging two rows
- Multiplying a row by a non-zero constant
- Adding a multiple of a row to another row

Row-echelon Form

A leading entry is the leftmost non-zero entry in a row of the matrix. A matrix is in row-echelon form if and only if

- For any row with a leading entry, all elements below that entry in the same column as it are zero.
- For any two rows, the leading entry of a lower row is further right than the leading entry of the higher row.
- Any row consisting solely of zeros is lower than any row with non-zero entries.

For example, of the below the first and second are in row-echelon form, while the third

$$\begin{bmatrix} 1 & -2 & 3 & 4 & 5 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 4 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{array}\right] \left[\begin{array}{cccc|c} 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 3 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & -3 & 6 & -4 & 9 \end{array}\right]$$

is not in row-echelon form. The first condition fails, as while the first row has a leading entry, non-zero entries follow in the same column. In addition, the rows are not correctly ordered by earliest leading entry.

Gaussian Elimination

Gaussian elimination is a recursive algorithm for putting matrices into row-echelon form.

- First switch rows around to bring a non-zero entry to the topmost leftmost position possible.
- After this, use suitable multiples of the first row to zero all entries in the same column.
- Repeat from step one, ignoring the first row.

Example

$$\begin{array}{r} x-3y+2z=11 \\ 2x-3y-2z=13 \\ 4x-2y+5z=31 \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & -3 & 2 & 11 \\ 2 & -3 & -2 & 13 \\ 4 & -2 & 5 & 31 \end{array}\right] \begin{array}{c} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 4R_1 \end{array} \to \left[\begin{array}{ccc|c} 1 & -3 & 2 & 11 \\ 0 & 3 & -6 & -9 \\ 0 & 10 & -3 & -13 \end{array}\right]$$

$$R_3 \to R_3 - \frac{10}{3}R_2 \to \left[\begin{array}{ccc|c} 1 & -3 & 2 & 11 \\ 0 & 3 & -6 & -9 \\ 0 & 0 & 17 & 17 \end{array}\right] \Rightarrow \begin{array}{c} x-3y+2z=11 \\ 0x+3y-6z=-9 \\ 0x+0y+17z=17 \end{array}$$

$$\Rightarrow z=1 \Rightarrow 3y-6=-9 \Rightarrow y=-1$$

$$y=-1 \,\&\, z=1 \Rightarrow x=6$$

While here we converted the matrix back to equations before solving, we could instead perform back substitution while still in matrix form. In this case, we are converting the matrix to *reduced row-echelon form*. A matrix in this form has the following properties

- It is in row-echelon form
- Each leading entry is equal to 1
- In each column containing a leading 1, all other entries are 0

Note that for a matrix, the reduced row-echelon form is unique.

Gauss-Jordan Elimination

Gauss-Jordan elimination is an algorithm to convert matrices to reduced row-echelon form.

- Use Gaussian elimination to reduce to row-echelon form.
- Multiply each non-zero row by an appropriate number to create a leading 1.
- Add row multiples to create zeros above leading entries.

Example

Continuing from the earlier example, where we found a row-echelon form for a linear sytem.

$$\left[\begin{array}{ccc|c} 1 & -3 & 2 & 11 \\ 0 & 3 & -6 & -9 \\ 0 & 0 & 17 & 17 \end{array}\right] \begin{array}{c} R_2 \to \frac{1}{3}R_2 \\ R_3 \to \frac{1}{17}R_3 \end{array} \left[\begin{array}{ccc|c} 1 & -3 & 2 & 11 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & 1 \end{array}\right] \begin{array}{c} R_2 \to R_2 + 2R_3 \\ R_1 \to R_1 - 2R_3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & -3 & 0 & 9 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right] R_1 \to R_1 + 3R_2 \left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \begin{array}{c} x = 6 \\ y = -1 \\ z = 1 \end{array}$$

In this case, it is clear that the reduced row echelon form is unique. With this, we have our general process for solving linear systems.

- First, create the augmented matrix from the equations.
- Use Gaussian elimination to get to row-echelon form.
- From here one can either directly back-substitute or use Gauss-Jordan elimination to get to reduced row-echelon form.
- At this stage, either the solutions will be evident, or it will be clear either that there are no or infinitely many solutions.

Solutions

- If a system of linear equations has no solutions, the system is inconsistent. In this case, row operations can be used to rearrange for a clear logical failure.
- If the system has exactly one or infinitely many solutions, it is consistent. Every homogenous system is consistent as their is the solution of zeroing all variables.

An inconsistent system can be identified by reducing to row-echelon form. In this case, if there is a row with a leading entry of the solution (i.e. all coefficients are zero and the solution is non-zero).

For a consistent system with one solution, there will be n non-zero rows for an n coefficient matrix. For a consistent system with an infinite number of solutions, there will be one or more rows of all zeros upon conversion to reduced row-echelon form.

In the case of an infinite number of solutions, the columns lacking a leading entry in will appear in some other equations. The values for these variables can then be chosen abritrarily, implying the values of other variables. Some examples

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 9 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & 5 & 1 \\ 0 & 0 & 1 & 0 & 6 & 2 \\ 0 & 0 & 0 & 1 & 7 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Here, in the left matrix the x_4 variable can be chosen as any value, and will set the values of the other columns. Likewise in the right matrix, x_2 and x_5 lack singly defined values.

In general, for a consistent system with n variables, there is one free variable for each variable without a leading entry.

Matrices

A matrix is a rectangular array of numbers with m rows and n columns.

$$A = 0_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad A_{1,2} = 3 \Rightarrow A = \begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Special Matrices

Some special matrix arrangements exist.

- A *square matrix* has the same number of rows and columns.
- A matrix with a single row is a *row matrix* or *row vector*

- A square matrix with $A_{ij} = 0$ for $i \neq j$ is called a diagonal matrix.
- A square matrix with $A_{ij} = 0$ for $i > j$ is called an *upper triangular matrix*.
- A square matrix with $A_{ij} = 0$ for $i < j$ is a *lower triangular matrix*.
- A matrix where all entries are 0 is a *zero matrix*, for a specific size denoted $0_{m,n}$.
- An identity matrix is a diagonal matrix with all non-zero fields equal to 1, denoted $I_{m,n}$.

Operations

Trace

The *trace* of a $n \times n$ matrix is given by

$$\mathrm{Tr}(A) = A_{11} + A_{22} + \ldots + A_{nn}$$

i.e. the sum of the diagonal elements.

Scalar Multiplication

Scalar multiplication of a matrix is performed by multiplying every entry by the relevant scalar.

$$2 \cdot \begin{bmatrix} 1 & 2 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 2 & -6 \end{bmatrix}$$

Naturally $1 \cdot A = A$ and $0 \cdot A = 0_{m,n}$.

Matrix Addition

For two matrices of the same size, we add entrywise to find the sum. To subtract we simply add $-1 \cdot B$.

$$A = \begin{bmatrix} 2 & 0 & -3 \\ 1 & -1 & 3 \end{bmatrix}, B = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$A + 3B = \begin{bmatrix} 2 & 0 & -3 \\ 1 & -1 & 3 \end{bmatrix} + \begin{bmatrix} -3 & -3 & 3 \\ 0 & 3 & 6 \end{bmatrix} = \begin{bmatrix} -1 & -3 & 0 \\ 1 & 2 & 9 \end{bmatrix}$$

Matrix addition has the same properties as addition of scalar addition including commutativity and associativity.

Matrix Multiplication

For two matrixes A of size $m \times n$ and B of size $n \times q$, we can perform multiplication as B has the same number of rows as A does columns. Unlike addition it is not performed elementwise. It is performed according to

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

This can be thought of as a dot product between the relevant row of A with the relevant column of B.

$$(AB)_{ij} = (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } B)$$

The output size will be the number of rows from the first matrix m and the number of columns in the second, q.

$$\begin{bmatrix} 1 & -1 \\ 3 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 0 \\ -7 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 + -1 \cdot -7 & 1 \cdot 0 + -1 \cdot 1 \\ 3 \cdot 4 + 0 \cdot -7 & 3 \cdot 0 + 0 \cdot 1 \\ 0 \cdot 4 + 1 \cdot -7 & 0 \cdot 0 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 11 & -1 \\ 12 & 0 \\ -7 & 1 \end{bmatrix}$$

Note that this process is not commutative. $AB \neq BA$. For example

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

$$AB = \begin{bmatrix} 1 \cdot 4 + 0 \cdot 2 & 1 \cdot 3 + 0 \cdot 1 \\ 2 \cdot 4 + 3 \cdot 2 & 2 \cdot 3 + 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 14 & 9 \end{bmatrix}$$

$$BA = \begin{bmatrix} 4 \cdot 1 + 3 \cdot 2 & 4 \cdot 0 + 3 \cdot 3 \\ 2 \cdot 1 + 1 \cdot 2 & 2 \cdot 0 + 1 \cdot 3 \end{bmatrix} = \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}$$

Matrix multiplication has the properties of

- Left distributivity, that $A(B + C) = AB + AC$.
- Right distributivity, that $(A + B)C = AC + BC$.
- Associative, that $A(BC) = ((AB)C)$.
- $A(\alpha B) = \alpha(AB)$
- $AI_n = I_m A = A$ where A has size $m \times n$.

A square matrix can be multiplied by itself to effect exponentiation.

Adjacency Matrices

An application of matrices is as a representation of a graph. We can use an $n \times n$ matrix to show the edges in a graph of n vertices.

$$\left[\begin{array}{c|ccccc} & v_1 & v_2 & v_3 & v_4 & v_5 \\ \hline v_1 & 0 & 1 & 0 & 0 & 1 \\ v_2 & 1 & 0 & 1 & 1 & 1 \\ v_3 & 0 & 1 & 0 & 1 & 0 \\ v_4 & 0 & 1 & 1 & 0 & 1 \\ v_5 & 1 & 1 & 0 & 1 & 0 \end{array}\right]$$

Here we have a 5 vertex graph, and we can tell from the matrix that for instance v_1 has edges to v_2 and v_5. A disconnected graph will simply be 0_n. The adjacency matrix is always symmetrical across the diagonal. If there are no edges looping back to their start, the diagonal will be zeroes.