Linear Algebra

Linear algebra originated in methods to solve systems of linear equations. Consider two equations in two unknowns, **x** & **y**, where **a, b, c, d, p** & **q** are known constants:

$$\mathbf{ax + by = p}$$
$$\mathbf{cx + dy = q} \quad .$$

Eliminate the variables one at a time in the usual way (multiply the first equation by **d**, the second by **b**, subtract to obtain: **adx – bcx = dp – bq**) and solve:

$$\mathbf{x} = \frac{\mathbf{dp - bq}}{\mathbf{ad - bc}} \quad \text{and} \quad \mathbf{y} = \frac{\mathbf{aq - cp}}{\mathbf{ad - bc}} \quad .$$

We may view the coefficients as a 2×2 matrix:

$$\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix} \quad \text{with determinant} \quad \mathbf{ad - bc} = \begin{vmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{vmatrix} \quad .$$

With this notation, the above solutions may be written:

$$\mathbf{x} = \frac{\begin{vmatrix} \mathbf{p} & \mathbf{b} \\ \mathbf{q} & \mathbf{d} \end{vmatrix}}{\begin{vmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{vmatrix}} \quad \text{and} \quad \mathbf{y} = \frac{\begin{vmatrix} \mathbf{a} & \mathbf{p} \\ \mathbf{c} & \mathbf{q} \end{vmatrix}}{\begin{vmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{vmatrix}} \quad .$$

Consider these three equations in three unknowns, **x, y & z:**

$$2\mathbf{x - y + z = 3}$$
$$\mathbf{-x + 3y - z = 2} \qquad \text{check that the solution is:}$$
$$\mathbf{x + 2y + z = 8} \qquad \mathbf{x = 1, y = 2, z = 3.}$$

One would solve it by adding multiples of one equation to another. If one replaces the system with the "augmented matrix" and does these row operations (r1=row#1):

$$\begin{pmatrix} \mathbf{2} & \mathbf{-1} & \mathbf{1} & \mathbf{3} \\ \mathbf{-1} & \mathbf{3} & \mathbf{-1} & \mathbf{2} \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & \mathbf{8} \end{pmatrix} \xrightarrow[\mathbf{(\frac{1}{5})r2;\ r1=r1-(2)r2}]{\mathbf{r1=r1+r2;\ r2=r2+r3;\ r3=r3-r1;}} \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{2} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{3} \end{pmatrix}$$

This result is called reduced row echelon form (rref). The solution is in the last column. The same steps operating on another "augmented matrix" yield this result:

$$\begin{pmatrix} \mathbf{2} & \mathbf{-1} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{-1} & \mathbf{3} & \mathbf{-1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{3/5} & \mathbf{-2/5} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1/5} & \mathbf{1/5} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} \end{pmatrix}$$

Matrix Concepts

Define an n×m **matrix** to be an array of numbers with n rows and m columns. A n×1 matrix is a **column vector** and a 1×n matrix is a **row vector**. As examples:

$\mathbf{A} = (a_{ij})$ **For** 3×4 $\mathbf{A} = \begin{pmatrix} 2 & 1 & -3 & 5 \\ -4 & 0 & -2 & 3 \\ -1 & -5 & 4 & 0 \end{pmatrix}$ then $a_{12} = \mathbf{1}$, $a_{23} = \mathbf{-2}$, $a_{34} = \mathbf{0}$.

$\mathbf{B} = (b_{ij})$ **For** 4×2 $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ -4 & 0 \\ -1 & -5 \\ 3 & -2 \end{pmatrix}$ then $b_{11} = \mathbf{2}$, $b_{32} = \mathbf{-5}$.

$\mathbf{X} = (x_i)$ **For** 4×1 $\mathbf{X} = \begin{pmatrix} 1 \\ 3 \\ -2 \\ 4 \end{pmatrix}$. $\mathbf{Y} = (y_j)$ **For** 1×4 $\mathbf{Y} = \begin{pmatrix} 3 & 2 & 1 & 5 \end{pmatrix}$.

The subscripted numbers are called the **components**. The number-pair of components are called the **dimensions**. A vector has one dimension. The **transpose** of a n×m matrix, $\mathbf{A=(a_{ij})}$, is the m×n marix, $\mathbf{A^T=(a_{ji})}$. A matrix may be multiplied by a real number, σ, called a **scalar** : $\sigma\mathbf{A} = (\sigma \cdot a_{ij})$. Two matrices of the same dimensions may be added: $\mathbf{A + B} = (a_{ij} + b_{ij})$.

A n×p matrix may be multiplied by a p×m matrix to yield a n×m matrix:

$\mathbf{A \bullet B} = \left(\sum_{k=1}^{p} a_{ik} \cdot b_{kj} \right)$. Important: $\mathbf{A \bullet B \neq B \bullet A}$; but $\mathbf{(A \bullet B) \bullet C = A \bullet (B \bullet C)}$ and $\mathbf{A \bullet (B + C) = A \bullet B + A \bullet C}$.

Using the above:

$$2\mathbf{A} = \begin{pmatrix} 4 & 2 & -6 & 10 \\ -8 & 0 & -4 & 6 \\ -2 & -10 & 8 & 0 \end{pmatrix} = \mathbf{A + A}.$$

Using the above: $\mathbf{A \bullet B} = \begin{pmatrix} 18 & 7 \\ 3 & 0 \\ 14 & -21 \end{pmatrix}$; where 1,2[th] entry is **7** = **2(1) + 1(0) + (-3)(-5) + 5(-2);** across 1[st] row, down 2[nd] column.

Using the above: $\mathbf{A \bullet X} = \begin{pmatrix} 31 \\ 12 \\ -24 \end{pmatrix}$ and $\mathbf{Y \bullet B} = \begin{pmatrix} -11 & -12 \end{pmatrix}$.

Square Matrices and Vectors

Using the definition of matrix multiplication, a square n×n matrix, **A** , defines a function or linear transformation on n×1 column vectors, **X & Y**:

$$\mathbf{Y} = \mathbf{A(X)} = \mathbf{A \bullet X} \ .$$

It is easy to verify that a linear transformation is linear, that is, it satisfies:

$$\mathbf{A}(\alpha\mathbf{X} + \beta\mathbf{Y}) = \alpha\mathbf{A}(\mathbf{X}) + \beta\mathbf{A}(\mathbf{Y}) \text{ , for scalars } \alpha \ \& \ \beta \ .$$

Composition of these matrix functions is the matrix which results from matrix multiplication: $(\mathbf{A} \circ \mathbf{B})(\mathbf{X}) = \mathbf{A}(\mathbf{B}(\mathbf{X})) = (\mathbf{A}\bullet\mathbf{B})\bullet\mathbf{X} = \mathbf{A}\bullet(\mathbf{B}\bullet\mathbf{X})$ **.**

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 5 \\ 3 & 2 & 1 \\ 4 & -4 & 3 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} -1 & 2 & 4 \\ 3 & 0 & 1 \\ -3 & 2 & 5 \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$$

$$\mathbf{B}\bullet\mathbf{X} = \begin{pmatrix} 9 \\ 13 \\ 7 \end{pmatrix} \qquad (\mathbf{A}\bullet\mathbf{B})\bullet\mathbf{X} = \begin{pmatrix} 5 \\ 60 \\ 5 \end{pmatrix}$$

The identity matrix, **I**, has 1's on the diagonal and 0's elsewhere:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

For all appropriate **A** & **X**, **I** satisfies:

$$\mathbf{I}\bullet\mathbf{A} = \mathbf{A} = \mathbf{A}\bullet\mathbf{I} \quad \& \quad \mathbf{I}\bullet\mathbf{X} = \mathbf{X} \ .$$

A square matrix, **A**, is **invertible** if it has an inverse matrix, $\mathbf{A}^{-1}$, with the property:

$$\mathbf{A}^{-1}\bullet\mathbf{A} = \mathbf{A}\bullet\mathbf{A}^{-1} = \mathbf{I} \ .$$

For example:

$$\begin{pmatrix} 2 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 3/5 & -2/5 \\ 0 & 1/5 & 1/5 \\ -1 & -1 & 1 \end{pmatrix}$$

Inverse matrices may be computed by row operations using the rref of the augmented matrix **[AI]** , as was demonstrated for the above matrix on the first page.

Inverse Matrices and Solving Systems

If one expands the definition of matrix multiplication:

$$\mathbf{A}\bullet\mathbf{X} = \mathbf{Y} \qquad \text{becomes} \qquad \begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = y_1 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = y_n \end{array}$$

a system of n linear equations. Suppose **Y** is known but **X** is not. To solve this system is to find $\mathbf{A}^{-1}$ since, applying $\mathbf{A}^{-1}$ to both sides we obtain: $\mathbf{X} = \mathbf{A}^{-1} \bullet \mathbf{Y}$ **.**

TI-83 and TI-84 Calculators Do Matrix Calculations.

To create or edit a matrix **A:**

2^{nd}-x^{-1} = 2^{nd}-matrix (choose Edit, choose **A)**;
Enter dimensions and #'s (2^{nd}-mode = 2^{nd}-quit).

To calculate determinants, **det(A):**

2^{nd}-matrix(choose Math, choose **det**); 2^{nd}-matrix(choose **A**).

To calculate transposes, $\mathbf{A}^{\mathbf{T}}$**:**

2^{nd}-matrix(choose Math, choose **T**); 2^{nd}-matrix(choose **A**).

To calculate inverses, $\mathbf{A}^{-1}$**:**

2^{nd}-matrix(choose **A**); **"**$\mathbf{x}^{-1}$**"** .

To calculate reduced row echelon form, **rref(A):**

2^{nd}-matrix(choose Math, choose **rref**); 2^{nd}-matrix(choose **A**).

To calculate **A + B (A i**s n×m & **B i**s n×m):

2^{nd}-matrix(choose **A**) **"+"** 2^{nd}-matrix(choose **B**).

To calculate **A•B (A i**s n×p & **B i**s p×m);

2^{nd}-matrix(choose **A**) **"×"** 2^{nd}-matrix(choose **B**).

To solve for **C** in **A × C = D :**

Create **A**(as n×n) and **D**(as n×k);
2^{nd}-matrix(choose **A**) **"**$\mathbf{x}^{-1}$**" "×"** 2^{nd}-matrix(choose **D**).
(The n×k matrix displayed is **C** . Usually k = 1.

Metric Structure on Vectors

Instead of viewing a vector as a column matrix of numbers, one may view it as the coordinates of a point in **n-dimensional** space. Picturing an **arrow** from the origin to the point, one may view the vector as having **length** and **direction** . These are defined using the **inner product**. Let:

$$\langle X,Y\rangle = X^T \bullet Y \text{, the inner or dot product, and, } \langle X,X\rangle = \| X \|^2 .$$

Call $\|X\|$ the **norm** or **magnitude** or **length** of **X** . The angle between vectors, θ , is given by the generalized law of cosines:

$$\langle X,Y\rangle = \|X\| \, \|Y\| \cos(\theta) .$$

If $\langle X,Y\rangle = 0$, **X&Y** are called orthogonal (or perpendicular, since $\cos(\pi/2) = 0$.)

In 2-dimensional space, two orthogonal vectors, **X&Y**, form a rectangle with area $= \|X\| \, \|Y\|$. In 3-dimensional space, three mutually orthogonal vectors, **X&Y&Z** , form a rectangular prism with volume $= \|X\| \, \|Y\| \, \|Z\|$. By analogy, we may define the "hyper-volume" of a "hyper-prism" formed by **n** mutually orthogonal vectors in **n**-dimensional space to be the product of their norms. We can compute the area, volume and hyper-volume of "slanted" parallelograms, parallelepipeds and hyper- parallelepipeds by the usual procedure of cutting and pasting tri-angular pieces. Call each of these things a **"box"** with a measure of **"size"**.

The linear transformation defined by a matrix, **A** , will take a "box" to a "box". The ratio of the size of the output box to the size of the input box is **det(A)** , called the **determinant** of A. The definition of **det(A)** in the 2×2 case is:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad ; \quad \det(A) = ad - bc .$$

Fact: **A** is invertible if and only if $\det(A) \neq 0$. Indeed, in the 2×2 case:

$$A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \frac{1}{\det(A)}$$

In n-dimensions, the formula is messy to define and messier to calculate.

Determinants and Solving Systems

As before, matrix multiplication defines a systems of equations:

$$\mathbf{A \bullet X = Y} \quad \text{becomes} \quad \begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = y_1 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = y_n \end{array}$$

To solve the system of linear equations is to find **X** for given **A** and **Y**. The unique solution, **X** , if there is one, has the coordinates given by:

$$x_k = \frac{\mathbf{det(A,k,Y)}}{\mathbf{det(A)}}$$; where **(A,k,Y)** is **A** with the **Y** in kth column.

Determinants and Linear Independence

Given **k** vectors of dimension **n**, $\mathbf{X^1, X^2, \ldots, X^k}$, (Superscripts are used to indicate different vectors; subscripts indicate the components of a vector.), a linear combination is a sum of the form:

$$\sigma_1\mathbf{X}^1 + \sigma_2\mathbf{X}^2 + \ldots + \sigma_k\mathbf{X}^k$$, where the σ_j are k scalars.

The k vectors are said to be linearly dependent if it is possible for a linear combination of them to be the zero vector with some of the scalars non-zero:

$$\sigma_1\mathbf{X}^1 + \sigma_2\mathbf{X}^2 + \ldots + \sigma_k\mathbf{X}^k = \mathbf{0}, \text{ with } \sigma_j \neq 0 \text{ for some of the j's.}$$

If not linear dependent, they are called linearly independent. If $\mathbf{k > n}$, then the vectors must be dependent. If $\mathbf{k \leq n}$, linearly independent vectors are said to span a space of dimension **k** (they form a **k**-dimensional "box"). A set of **n** linearly independent vectors of dimension **n** is called a fundamental set or basis. One may arrange a basis to form the $\mathbf{n \times n}$ matrix: $(\mathbf{X^j}) = ((x_i)^{\mathbf{j}})$ **.**

Proposition: $\{\mathbf{X^j}\}$, **n** vectors of dimension **n**, are linearly independent, that is, they are a basis if and only if $\mathbf{det}((x_i)^{\mathbf{j}}) \neq \mathbf{0}$.

Given a basis, any vector may be written as a unique linearly combination of the vectors in the basis. Precisely, given a basis ,$\{\mathbf{X^j}\}$, of vectors and any vector, **Y**, of the same dimension **n**, there are **n** unique scalars, $\{\sigma_j\}$, such that:

$$\mathbf{Y} = \sigma_1\mathbf{X}^1 + \sigma_2\mathbf{X}^2 + \ldots + \sigma_n\mathbf{X}^n \text{ .}$$

Eigenvalues and Eigenvectors

An **n×n** matrix, **A ,** is not-invertible if and only if it is **singular**, that is, there is a vector, **X** , with **A•X = 0** .

Remark: This means that **A** will collapse the hyper-size of a hyper-box to a lower dimensional box with **n**-dimensional hyper-size = zero.

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ -2 & -4 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 & 0 \\ 1 & 0 & 1 & 0 \\ -2 & -4 & -6 & 0 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

There is no unique solution. The determinant is:

$$\mathbf{det} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ -2 & -4 & -6 \end{pmatrix} = \mathbf{0}$$

A solution must have: y + z = 0 & x + z = 0.

So, any multiple of: $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ is a solution, an eigenvector for eigenvalue = 0.

An **n×n** matrix has an **eigenvalue** **λ** with **eigenvector** **X** if: **A•X = λX** . This means that the matrix (linear transformation) **A - λI** is singular – and hence:

det (A - λI) = 0 , which yields a polynomial of degree **n** in **λ** .

There will be **n** values of **λ** which satisfy this requirement. Those values of **λ** may be real or complex, and must be counted with multiplicity (Fundamental Theorem of Algebra). For a symmetric **A,** $\mathbf{A^T=A}$**,** all eigenvalues are real.

For: $\begin{pmatrix} 3 & -1 \\ 4 & -2 \end{pmatrix}$ **λ** = 2, -1 **X**'s are: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$

For: $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ **λ** = 2, -1, -1 **X**'s are: $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$

For: $\begin{pmatrix} -1/2 & 1 \\ -1 & -1/2 \end{pmatrix}$ **λ** = $-1/2 \pm i$ **X**'s are: $\begin{pmatrix} 1 \\ i \end{pmatrix}$ $\begin{pmatrix} 1 \\ -i \end{pmatrix}$