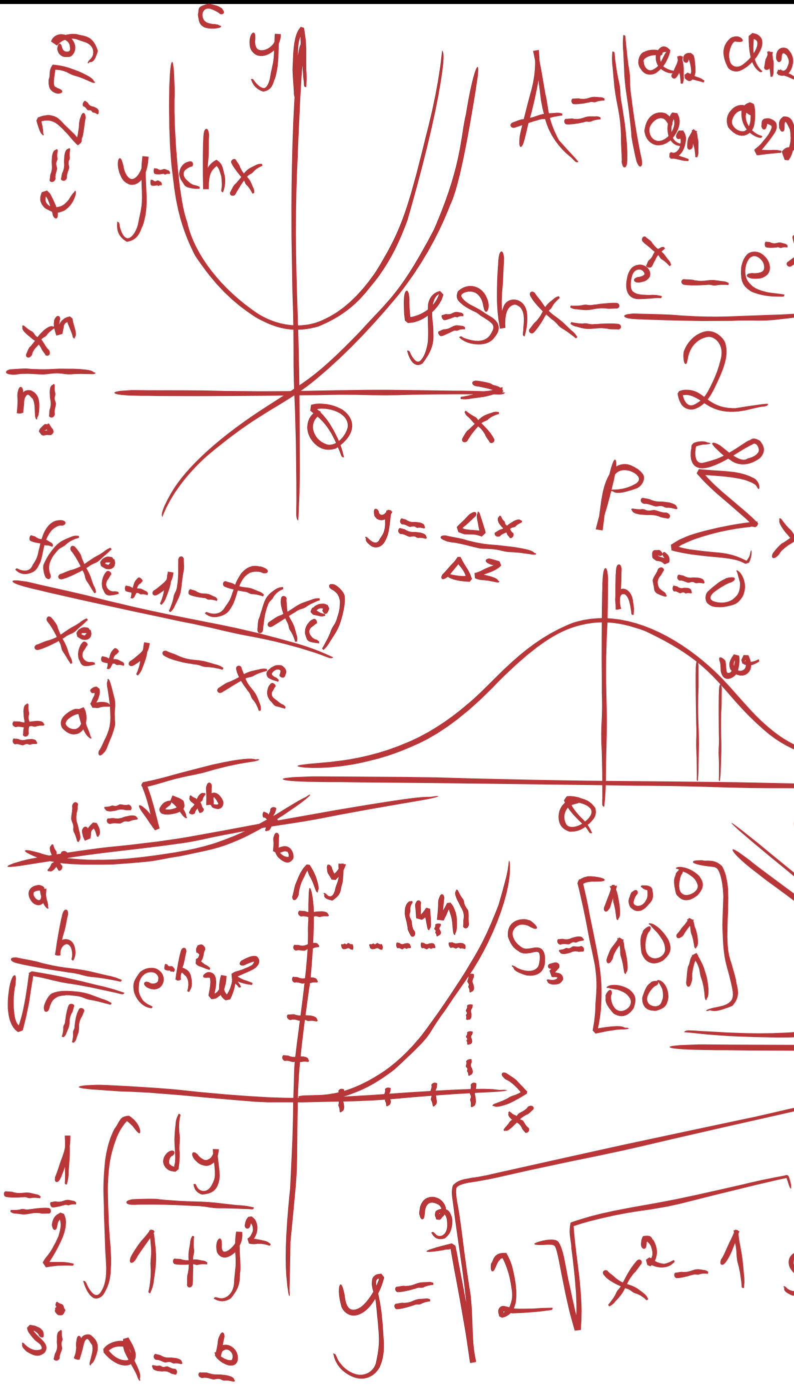
THE ARTISTS OF DATA SCIENCE PRESENT...

THE MATH YOU ABSOLUTELY NEED FOR DATA SCIENCE



CURATED BY:

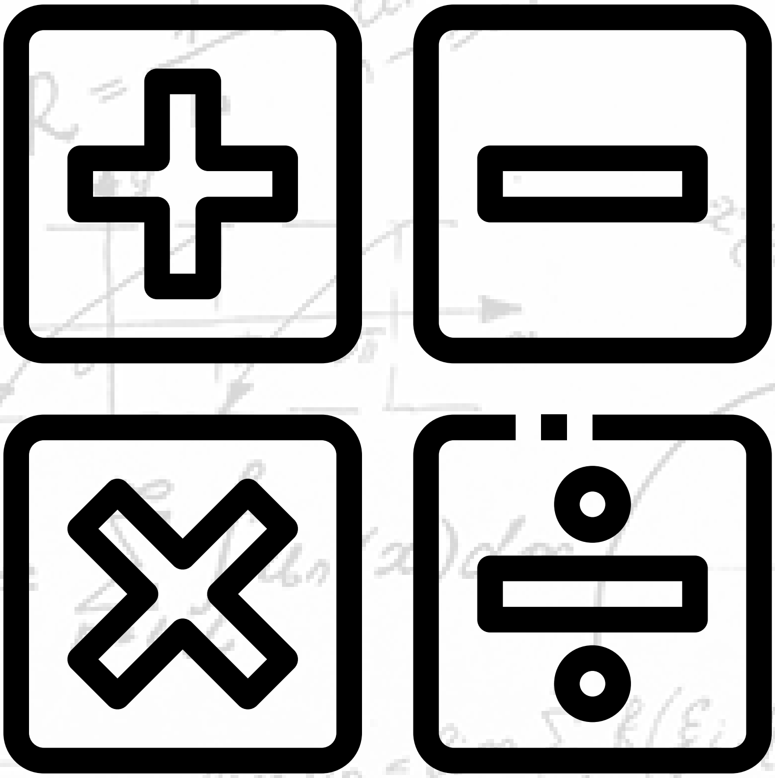
Harpreet Sahota

The Artists of Data Science

OVERVIEW

Yes, you need to know math to be in data science. But do you need to know all of it to get your first job in data science? **Absolutely not.** What you do need is a firm grasp on some fundamental concepts. If you understand these fundamentals then learning anything else you need won't be too much of a stress. Focus on the fundamentals, and know them intuitively.

WHAT I BELIEVE



I believe that any body can learn the essential tools of mathematics for data science. Just because you didn't major in math or stats or computer science doesn't mean that you can't learn those skills. At one point in my life I didn't know how to walk. Now I walk all the time. There was also a point when I didn't know how to code, and now I do it daily. You can learn anything.

OBJECTIVES

THE GOAL OF THIS PAPER

I'll start with the basics: Listing out everything I feel you need to know before even thinking about applying for your first job in data science. I'll link to my favorite resources so you can go more in depth. These are ones that I have hand picked for their clarity and delivery of content.

KEY POINTS

- This list is not exhaustive. There's gonna be some pedantic academic out there talking out of mouth about:"Well you need to know this, and that, and what about that". *If you're that person, relax. Close this document now. And go on about your day.*
- The skills outlined here are the foundations. They are the fundamentals on which everything else is built. If you study these, know them well, understand them intuitively, then you'll have no problem picking up more complex concepts.
- This is not a textbook. I will generously link to resources that explain things far better than I can. They're not random resources. They are ones that I fully co-sign.
- Pretty much everything here is lifted from various university courses. Mostly from Penn State. They're awesome.

ALGEBRA

Knowledge of the following mathematical operations is required for data science:

- **Addition**
- **Subtraction**
- **Division**
- **Multiplication**
- **Radicals (i.e., square roots)**
- **Exponents**
- **Summations (∑)**
- **Factorials (!)**
- **Scientific Notation**

When performing a series of mathematical operations, begin with those inside parentheses or brackets. Next, calculate any exponents or square roots. This is followed by multiplication and division, and finally, addition and subtraction.

- **Parentheses**
- **Exponents & Square Roots**
- **Multiplication and Division**
- **Addition and Subtraction**

Simplify: $(5 + \frac{9}{3})^2$

Answer

$$
\begin{aligned}
(5 + \frac{9}{3})^2 &= (5+3)^2 && \text{(Parentheses first, 9/3)} \\
&= 8^2 && \text{(Add inside parentheses)} \\
&= 64 && \text{(Square 8)}
\end{aligned}
$$

Simplify: $\frac{2^2 + 3^2 + 4^2}{3 - 1}$

Answer

$$
\begin{aligned}
\frac{2^2 + 3^2 + 4^2}{3 \quad 1} &= \frac{4 + 9 + 16}{2} && \text{(Exponents in the numerator)} \\
&= \frac{29}{2} && \text{(Simplify the numerator)} \\
&= 14.5 && \text{(Divide)}
\end{aligned}
$$

Take a self assessment here and check your results here.

The Artists of Data Science
with Harpreet Sahota

CALCULUS

For long-term success in the field of data science, it is imperative that you have a working knowledge of multidimensional calculus. We'll do a simple review of the calculus techniques most frequently used in white papers and research papers for statistics and machine learning:

- **Summation and Series**
- **Limits**
- **Differentiation**
- **Integration**
- **Multivariate calculus**

Summations

First, it is important to review the notation. The symbol, $\sum$, is a summation. Suppose we have the sequence, $a_1, a_2, \cdots, a_n$, denoted $\{a_n\}$, and we want to sum all their values. This can be written as

$$\sum_{i=1}^{n} a_i$$

Here are some special sums:

1. $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$
2. $\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$
3. The Binomial Theorem:

It is possible to expand any power of $x + y$ to the sum

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$$

where

$$\binom{n}{i} = \frac{n(n-1)(n-2)\cdots(n-i-1)}{i!} = \frac{n!}{(n-i)!i!}$$

Learn more about The Binomial Theorem

The Artists of Data Science with Harpreet Sahota

CALCULUS - SERIES

Series

When n is a finite number, the value of the sum can be easily determined. How do we find the sum when the sequence is infinite? For example, suppose we have an infinite sequence, $a_1, a_2, \cdots$. The infinite series is denoted:

$$S = \sum_{i=1}^{\infty} a_i$$

For infinite series, we consider the partial sums. Some partial sums are

$$S_1 = \sum_{i=1}^{1} a_i = a_1$$

$$S_2 = \sum_{i=1}^{2} a_i = a_1 + a_2$$

$$S_3 = \sum_{i=1}^{3} a_i = a_1 + a_2 + a_3$$

$$\vdots$$

$$S_n = \sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n$$

An infinite series **converges** and has sum S if the sequence of partial sums, $\{S_n\}$ converges to S. Thus, if

$$S = \lim_{n \to \infty} \{S_n\}$$

then the series converges to S. If $\{S_n\}$ diverges, then the series diverges.

Learn more about convergence and divergence of series' here

CALCULUS - SERIES

Geometric series

A geometric series has the form

$$S = \sum_{k=1}^{\infty} ar^{k-1} = a + ar + ar^2 + ar^3 + \cdots$$

where $a \neq 0$. A geometric series converges to $\frac{a}{1-r}$ if $|r| < 1$, but diverges if $|r| \geq 1$.

A special case of the geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$$

for $-1 < x < 1$.

The Taylor (or Maclaurin) series of e^x:

The series:

$$\sum_{i=0}^{\infty} \frac{x^i}{i!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

for $-1 \leq x \leq 1$ converges to e^x.

- **Learn more about geometric series here**
- **Learn more about the Taylor series here**

CALCULUS - DERIVATIVES

The definition of a derivative is

$$f'(x) = \frac{d}{dx} f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

The derivative is the slope of the tangent line to the graph of f(x), assuming the tangent line exists.

Here are some common rules you should know:

- **Power Rule**
- **Product Rule**
- **Quotient Rule**
- **Chain Rule**
- **L'Hopital's Rule**

Learn more about what the limit of a function is here and here.

Learn more about derivaties here.

CALCULUS - INTEGRALS

For a function f(x), its definite integral is:

$$\int f(x)\, dx = F(x) + C, \qquad \text{where } F'(x) = f(x)$$

Common Integrals and Rules

a. $\int_a^a f(x)dx = 0$

b. $\int_a^b f(x)d(x) = -\int_b^a f(x)d(x)$

c. $\int x^r dx = \frac{x^{r+1}}{r+1} + C$

Some other concepts related to integrals you should be aware of:

- **The Fundamental Theorem of Calculus**
- **Integration using substitution**
- **Integration by parts**

Learn more on integrals here, and here.

CALCULUS - MULTIVARIATE

Partial Derivatives

Let's begin with **Partial Derivatives**. Suppose we have the function $f(x, y)$. The partial derivative with respect to x would be

$$f_x(x, y) = \lim_{h \to 0} \frac{f(x-h, y)}{x-h}$$

Similarly, the partial derivative of $f(x, y)$ with respect to y would be

$$f_y(x, y) = \lim_{h \to 0} \frac{f(x, y-h)}{y-h}$$

The notation for partial derivatives is not the same for all texts. You should be able to recognize the different forms. The notation, for example, for the partial derivative of $f(x, y)$, with respect to x, could be denoted as:

$$f_x(x, y) = \frac{\partial}{\partial x} f(x, y) = \frac{\partial f}{\partial x}$$

Learn more about partial derivatives here and here.

The Artists of Data Science with Harpreet Sahota

CALCULUS - MULTIVARIATE

Integrating over regions is important in calculus and has many applications in probability theory. Suppose we have a function *f(x,y)* over some region *R*

$$\int\int_R f(x,y)\ dxdy$$

Consider the rectangular region defined by $a \leq x \leq b$ and $c \leq y \leq d$, or $R = [a,b] \times [c,d]$. Then the iterated integral would be:

$$\int_c^d \left[\int_a^b f(x,y)dx\right]dy = \int_a^b \left[\int_c^d f(x,y)dy\right]dx$$

When the region is not rectangular, things can get complicated. It is important to draw out the support space and consider the region when building these double integrals.

Learn more about double integrals here and here.

Visit here to learn how integrals are used in probability theory.

The Artists of Data Science
with Harpreet Sahota

LINEAR ALGEBRA – MATRICES

Matrix

A **matrix** is a rectangular collection of numbers. Generally, matrices are denoted as bold capital letters. For example:

$$A = \begin{pmatrix} 1 & -5 & 4 \\ 2 & 5 & 3 \end{pmatrix}$$

A is a matrix with two rows and three columns. For that reason, it is called a 2 by 3 matrix. This is called the **dimension** of a matrix.

Dimension

The **dimension** of a matrix is expressed as number of rows × number of columns. So,

$$B = \begin{pmatrix} 1 & -5 & 4 \\ 5 & 3 & -8 \\ 1 & 5 & 4 \\ 2 & 5 & 3 \end{pmatrix}$$

B is a 4 × 3 matrix. It is common to refer to elements in a matrix by subscripts, like so.

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \\ b_{4,1} & b_{4,2} & b_{4,3} \end{pmatrix}$$

With the row first and the column second. So in this case, $b_{2,1} = 5$ and $b_{1,3} = 4$.

The Artists of Data Science with Harpreet Sahota

LINEAR ALGEBRA – MATRICES

Vector

A **vector** is a matrix with only one row (called a **row vector**) or only one column (called a **column vector**). For example:

$$C = \begin{pmatrix} 2 & 7 & -3 & 5 \end{pmatrix}$$

C is a 4 dimensional row vector.

$$D = \begin{pmatrix} 2 \\ 9 \\ -3 \\ 3 \\ 6 \end{pmatrix}$$

D is a 5 dimensional column vector. An "ordinary" number can be thought of as a 1 × 1 matrix, also known as a **scalar**. Some examples of scalars are shown below:

$$E = \pi$$

$$F = 6$$

The Artists of Data Science with Harpreet Sahota

LINEAR ALGEBRA - MATRICES

Transpose a Matrix

To take the **transpose** of a matrix, simply switch the rows and column of a matrix. The transpose of A can be denoted as A' or A^T.

> For example
>
> $$A = \begin{pmatrix} 1 & -5 & 4 \\ 2 & 5 & 3 \end{pmatrix}$$
>
> $$A' = A^T = \begin{pmatrix} 1 & 2 \\ -5 & 5 \\ 4 & 3 \end{pmatrix}$$

If a matrix is its own transpose, then that matrix is said to be **symmetric**. Symmetric matrices must be square matrices, with the same number of rows and columns.

> One example of a symmetric matrix is shown below:
>
> $$A = \begin{pmatrix} 1 & -5 & 4 \\ -5 & 7 & 3 \\ 4 & 3 & 3 \end{pmatrix} = A' = A^T$$

Learn more about the transpose of a matrix here.

The Artists of Data Science with Harpreet Sahota

LINEAR ALGEBRA - MATRICES

Matrix Addition

To perform matrix **addition**, two matrices must have the same dimensions. This means they must have the same number of rows and columns. In that case simply add each individual components, like below.

> For example
>
> $$A + B = \begin{pmatrix} 1 & -5 & 4 \\ 2 & 5 & 3 \end{pmatrix} + \begin{pmatrix} 8 & -3 & -4 \\ 4 & -2 & 9 \end{pmatrix} = \begin{pmatrix} 1+8 & -5-3 & 4-4 \\ 2+4 & 5-2 & 3+9 \end{pmatrix} = \begin{pmatrix} 9 & -8 & 0 \\ 6 & 3 & 12 \end{pmatrix}$$

Matrix addition does have many of the same properties as "normal" addition.

$$A + B = B + A$$

$$A + (B + C) = (A + B) + C$$

In addition, if one wishes to take the transpose of the sum of two matrices, then

$$A^T + B^T = (A + B)^T$$

Learn more about the matrix addition <u>here</u>.

LINEAR ALGEBRA - MATRICES

Matrix Scalar Multiplication

To multiply a matrix by a scalar, also known as **scalar multiplication**, multiply every element in the matrix by the scalar.

For example...

$$6 * A = 6 * \begin{pmatrix} 1 & -5 & 4 \\ 2 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 6*1 & 6*-5 & 6*4 \\ 6*2 & 6*5 & 6*3 \end{pmatrix} = \begin{pmatrix} 6 & -30 & 24 \\ 12 & 30 & 18 \end{pmatrix}$$

To multiply two vectors with the same length together is to take the **dot product**, also called **inner product**. This is done by multiplying every entry in the two vectors together and then adding all the products up.

For example, for vectors *x* and *y*, the dot product is calculated below

$$x \cdot y = \begin{pmatrix} 1 & -5 & 4 \end{pmatrix} * \begin{pmatrix} 4 & -2 & 5 \end{pmatrix} = 1*4 + (-5)*(-2) + 4*5 = 4 + 10 + 20 = 34$$

Learn more about the matrix scalar multiplication <u>here</u>.

The Artists of Data Science
with Harpreet Sahota

LINEAR ALGEBRA - MATRICES

Matrix Multiplication

To perform **matrix multiplication**, the first matrix must have the same number of columns as the second matrix has rows. The number of rows of the resulting matrix equals the number of rows of the first matrix, and the number of columns of the resulting matrix equals the number of columns of the second matrix. So a 3 × 5 matrix could be multiplied by a 5 × 7 matrix, forming a 3 × 7 matrix, but one cannot multiply a 2 × 8 matrix with a 4 × 2 matrix. To find the entries in the resulting matrix, simply take the dot product of the corresponding row of the first matrix and the corresponding column of the second matrix.

> For example
>
> $$C * D = \begin{pmatrix} 3 & -9 & -8 \\ 2 & 4 & 3 \end{pmatrix} * \begin{pmatrix} 7 & -3 \\ 2 & 3 \\ 6 & 2 \end{pmatrix}$$
>
> $$C * D = \begin{pmatrix} 3*7+(-9)*(-2)+(-8)*6 & 3*(-3)+(-9)*3+(-8)*2 \\ 2*7+4*(-2)+3*6 & 2*(-3)+4*3+3*2 \end{pmatrix}$$
>
> $$C * D = \begin{pmatrix} 21+18-48 & -9-27-16 \\ 14-8+18 & -6+12+6 \end{pmatrix} = \begin{pmatrix} -9 & -52 \\ 24 & 12 \end{pmatrix}$$

Matrix multiplication has some of the same properties as "normal" multiplication , such as

$$A(BC) = (AB)C$$

$$A(B + C) = AB + AC$$

$$(A + B)C = AC + BC$$

However matrix multiplication is not communicative. That is to say *A**B* does not necessarily equal *B**A*. In fact, *B**A* often has no meaning since the dimensions rarely match up. However, you can take the transpose of matrix multiplication. In that case $(AB)^T = B^T A^T$.

Learn more about the matrix multiplication here.

LINEAR ALGEBRA - MATRIX PROPERTIES

Identity Matrices

An **identity matrix** is a square matrix where every diagonal entry is 1 and all the other entries are 0. The following two matrices are both identity matrices and diagonal matrices.

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

They are called identity matrices, because any matrix multiplied with an identify matrix equals itself. The diagonal entries of a matrix are the entries where the column and row number are the same. $a_{2,2}$ is a diagonal entry but $a_{3,5}$ is not. The **trace** of a *n* × *n* matrix is the sum of all the diagonal entries. In other words, for *n* × *n* matrix A, $trace(A) = tr(A) = \sum_{i=1}^{n} a_{i,i}$ For example:

$$trace(F) = tr(F) = tr \begin{pmatrix} 1 & 3 & 3 \\ 0 & 6 & 7 \\ -5 & 0 & 1 \end{pmatrix} = 1 + 6 + 1 = 8$$

The trace has some useful properties, namely that for same size square matrices *A* and *B* and scalar *c*,

$$tr(A) = tr(A^T)\ tr(A + B) = tr(B + A) = tr(A) + tr(B)\ tr(AB) = tr(BA)\ tr(cA) = c * tr(A)$$

Learn more about the identity matrix [here](#).

LINEAR ALGEBRA – MATRIX PROPERTIES

Determintants

The **determinate** of a square, 2 × 2 matrix A is

$$det(A) = |A| = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1} * a_{2,2} - a_{1,2} * a_{2,1}$$

For example

$$det(A) = |A| = \begin{vmatrix} 5 & 2 \\ 7 & 2 \end{vmatrix} = 5 * 2 - 2 * 7 = -4$$

In general, to find the determinate of a n × n matrix, choose a row or column like column 1, and take the determinates of the "minor" matrices inside the original matrix, like so:

$$det(C) = |C| = \begin{vmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{.,2} & \cdots & c_{n,n} \end{vmatrix}$$

$$det(C) = (-1)^{1+1} c_{1,1} \begin{vmatrix} c_{2,2} & \cdots & c_{2,n} \\ \vdots & \ddots & \vdots \\ c_{n,2} & \cdots & c_{n,n} \end{vmatrix} + (-1)^{2+1} c_{2,1} \begin{vmatrix} c_{1,2} & \cdots & c_{1,n} \\ c_{3,2} & \cdots & c_{3,n} \\ \vdots & \ddots & \vdots \\ c_{n,2} & \cdots & c_{n,n} \end{vmatrix} + \ldots$$

$$\ldots + (-1)^{n+1} c_{n,1} \begin{vmatrix} c_{1,2} & \cdots & c_{1,n} \\ \vdots & \ddots & \vdots \\ c_{n-1,2} & \cdots & c_{n-1,n} \end{vmatrix}$$

This is known as Laplace's formula,

Laplace's Formula

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} det(A_{-i,-j}) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} det(A_{-i,-j})$$

For any i, j, where $A_{-i,-j}$ is matrix A with row i and column j removed. This formula works whether one goes by rows, using the first formulation, or by columns, using the second formulation. It is easiest to use Laplace's formula when one chooses the row or column with the most zeroes.

LINEAR ALGEBRA - MATRIX PROPERTIES

Matrix Determinant Properties

The matrix determinate has some interesting properties.

$$det(I) = 1$$

where I is the identity matrix.

$$det(A) = det(A^T)$$

If A and B are square matrices with the same dimensions, then

$det(AB) = det(A) * det(B)$ and if A is a $n \times n$ square matrix and c is a scalar, then

$$det(cA) = c^n det(A)$$

Learn more about the identity matrix determinants here.

LINEAR ALGEBRA - MATRIX INVERSES

Inverse of a Matrix

The matrix B is the **inverse** of matrix A if $AB = BA = I$. This is often denoted as $B = A^{\ 1}$ or $A = B^{\ 1}$. When taking the inverse of the product of two matrices A and B,

$$(AB)^{\ 1} = B^{\ 1}A^{\ 1}$$

When taking the determinate of the inverse of the matrix A,

$$det(A^{\ 1}) = \frac{1}{det(A)} = det(A)^{\ 1}$$

Note that not all matrices have inverses. For a matrix A to have an inverse, that is to say for A to be **invertible**, A must be a square matrix and $det(A) \neq 0$. For that reason, invertible matrices are also called **nonsingular** matrices.

Learn more about matrix inversion here, here, and here.

LINEAR ALGEBRA - ADVANCED PROPERTIES

Orthogonal Vectors

Two vectors, *x* and *y*, are **orthogonal** if their dot product is zero.

> For example
>
> $$e \cdot f = (2 \quad 5 \quad 4) * \begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix} = 2 * 4 + (5) * (\quad 2) + 4 * 5 = 8 \quad 10 + 20 = 18$$
>
> Vectors e and *f* are not orthogonal.
>
> $$g \cdot h = (2 \quad 3 \quad -2) * \begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix} = 2 * 4 + (3) * (-2) + (-2) * 1 = 8 - 6 - 2 = 0$$

However, vectors *g* and *h* are orthogonal. Orthogonal can be thought of as an expansion of perpendicular for higher dimensions. Let $x_1, x_2, \ldots, x_n$ be *m*-dimensional vectors. Then a **linear combination** of $x_1, x_2, \ldots, x_n$ is any *m*-dimensional vector that can be expressed as

So *y* is a linear combination of x_1 and x_2. The set of all linear combinations of $x_1, x_2, \ldots, x_n$ is called the **span** of $x_1, x_2, \ldots, x_n$. In other words

$$span(\{x_1, x_2, \ldots, x_n\}) = \{v | v = \sum_{i=1}^{n} c_i x_i, c_i \in \mathbb{R}\}$$

A set of vectors $x_1, x_2, \ldots, x_n$ is **linearly independent** if none of the vectors in the set can be expressed as a linear combination of the other vectors. Another way to think of this is a set of vectors $x_1, x_2, \ldots, x_n$ are linearly independent if the only solution to the below equation is to have $c_1 = c_2 = \ldots = c_n = 0$, where $c_1, c_2, \ldots, c_n$ are scalars, and 0 is the zero vector (the vector where every entry is 0).

$$c_1 x_1 + c_2 x_2 + \ldots + c_n x_n = 0$$

If a set of vectors is not linearly independent, then they are called **linearly dependent**.

Learn more here.

The Artists of Data Science with Harpreet Sahota

LINEAR ALGEBRA - ADVANCED PROPERTIES

Norm of a vector or matrix

The **norm** of a vector or matrix is a measure of the "length" of said vector or matrix. For a vector *x*, the most common norm is the $\mathbf{L_2}$ **norm**, or **Euclidean norm**. It is defined as

$$||x|| = ||x||_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Other common vector norms include the $\mathbf{L_1}$ **norm**, also called the **Manhattan norm** and **Taxicab norm**.

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

Other common vector norms include the **Maximum norm**, also called the **Infinity norm**.

$$||x||_\infty = max(|x_1|, |x_2|, \ldots, |x_n|)$$

The most commonly used matrix norm is the **Frobenius norm**. For a $m \times n$ matrix A, the Frobenius norm is defined as:

$$||A|| = ||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} x_{i,j}^2}$$

- **Learn more about Euclidean norm here.**
- **Learn more about norms here.**

The Artists of Data Science with Harpreet Sahota

LINEAR ALGEBRA – ADVANCED PROPERTIES

Quadratic Form of a Vector

The **quadratic form** of the vector *x* associated with matrix *A* is

$$x^T Ax = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} x_i x_j$$

A matrix *A* is **Positive Definite** if for any non-zero vector *x*, the quadratic form of *x* and *A* is strictly positive. In other words, $x^T Ax > 0$ for all nonzero *x*.

A matrix *A* is **Positive Semi-Definite** or **Non-negative Definite** if for any non-zero vector *x*, the quadratic form of *x* and *A* is non-negative . In other words, $x^T Ax \geq 0$ for all non-zero *x*. Similarly,

A matrix *A* is **Negative Definite** if for any non-zero vector *x*, $x^T Ax < 0$. A matrix *A* is **Negative Semi-Definite** or **Non-positive Definite** if for any non-zero vector *x*, $x^T Ax \leq 0$.

LINEAR ALGEBRA - RANGE OF A MATRIX

Range of a matrix

The **range** of $m \times n$ matrix A, is the span of the n columns of A. In other words, for

$$A = [a_1 a_2 a_3 \dots a_n]$$

where $a_1, a_2, a_3, \dots, a_n$ are m-dimensional vectors,

$$range(A) = R(A) = span(\{a_1, a_2, \dots, a_n\}) = \{v | v = \sum_{i=1}^{n} c_i a_i, c_i \in \mathbb{R}\}$$

The dimension (number of linear independent columns) of the range of A is called the **rank** of A. So if 6×3 dimensional matrix B has a 2 dimensional range, then $rank(A) = 2$.

For example

$$C = \begin{pmatrix} 1 & 4 & 1 \\ 8 & 2 & 3 \\ 8 & 2 & -2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

C has a rank of 3, because x_1, x_2 and x_3 are linearly independent.

LINEAR ALGEBRA - NULLSPACE OF A MATRIX

Nullspace

p>The **nullspace** of a $m \times n$ matrix is the set of all n-dimensional vectors that equal the n-dimensional zero vector (the vector where every entry is 0) when multiplied by A. This is often denoted as

$$N(A) = \{v | Av = 0\}$$

The dimension of the nullspace of A is called the **nullity** of A. So if 6×3 dimensional matrix B has a 1 dimensional range, then $nullity(A) = 1$.

The range and nullspace of a matrix are closely related. In particular, for $m \times n$ matrix A,

$$\{w | w = u + v, u \in R(A^T), v \in N(A)\} = \mathbb{R}^n$$

$$R(A^T) \cap N(A) = \phi$$

This leads to the **rank--nullity theorem**, which says that the rank and the nullity of a matrix sum together to the number of columns of the matrix. To put it into symbols:

Rank--nullity Theorem

$$A \in \mathbb{R}^{m \times n} \Rightarrow rank(A) + nullity(A) = n$$

For example, if B is a 4×3 matrix and $rank(B) = 2$, then from the rank--nullity theorem, on can deduce that

$$rank(B) + nullity(B) = 2 + nullity(B) = 3 \Rightarrow nullity(B) = 1$$

Learn more about null space here.

The Artists of Data Science
with Harpreet Sahota

LINEAR ALGEBRA – OTHER TOPICS

- **Some other important topic include:**
- **Gauss-Jordan Elimination**
- **Using Gauss-Jordan to Solve a System of Three Linear Equations - Example 1**
- **Algebra - Matrices - Gauss Jordan Method Part 1 Augmented Matrix**
- **Gaussian Elimination**

The Artists of Data Science
with Harpreet Sahota

LINEAR ALGEBRA – EIGENDECOMPOSITION

Eigenvector of a matrix

An **eigenvector** of a matrix *A* is a vector whose product when multiplied by the matrix is a scalar multiple of itself. The corresponding multiplier is often denoted as $lambda$ and referred to as an **eigenvalue**. In other words, if *A* is a matrix, *v* is a eigenvector of *A*, and λ is the corresponding eigenvalue, then $Av = \lambda v$.

- **Learn about everything related to Eigens here.**
- **A concept especially important in PCA is the Singular Value Decomposition. Learn more about that here.**

The Artists of Data Science
with Harpreet Sahota

ADDITIONAL RESOURCES

All of the above can be reviewed below. This is a list of some of my personal favorite places that have helped me on my journey:

- **PatrickJMT**
- **NancyPi**
- **Krista King**
- **CatBug88**
- **MathByFives**

CLOSING REMARKS

There is a lot of math to learn. That is for sure.

But don't feel like you have to memorize everything. Go through the topics as quickly as you can get gain an awareness of the concepts and use this document as a refresher for when you forget some stuff.

You won't have to solve problems by hand on the job. But you will need to read papers, mostly academic ones.

So it's a good idea to be aware of these concepts so you can progress in your career.

Thanks!