Linear Algebra

Terry Yu

2023

Math 33A with Prof. Pablo

Contents

§1 Linear Equations and Matrices

Introduction to Linear Systems, Matrices, Vectors, and Gauss-Jordan Elimination

Linear Algebra is the study of linear equations and linear transformations.

Definition — A linear equation is has the form

$$ax_1 + bx_2 + \cdots + a_n x_n + b = 0$$

Given a system of linear equations, such as

$$\begin{cases} x + 2y + 3z = 39 \\ x + 3y + 2z = 34 \\ 3x + 2y + z = 26 \end{cases}$$

we can solve for the solution, or the value(s) of x, y, z which satisfies this equation. One way of solving it is with the *elimination method*, which involves a combination of scaling, adding/subtracting, and swapping the rows of equations.

Solving this equation yields

$$\begin{cases} x = 2.75 \\ y = 4.25 \\ z = 9.25 \end{cases}$$

Not only does this mean that those values of x, y, and z satisfy all 3 equalities, it also shows, geometrically, the point of intersection of the 3 equations when graphed on a 3d vector space.

Linear Algebra produces an alternate method of denoting and solving systems of linear equations, however, using **matrices**.

Matrix — A matrix is a rectangular array of numbers.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Shown above is a $n \times m$ size matrix (n rows and m columns).

There are a couple special matrixes, including the *zero matrix*,

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

or the $n \times n$ *unit/identity matrix*,

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Vector — Most importantly, a matrix with only one column, is called a *column vector*, or simply just a *vector*.

$$\vec{\mathbf{v}} = \begin{pmatrix} x \\ y \\ \vdots \end{pmatrix}$$

The set of all column vectors with n components is denoted by $\mathbb{R}^n$, known as the *vector space*.

A system of linear equations can be expressed through the use of an *augumented matrix*; shown below is a system of equations, and the corresponding augumented matrix.

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$$

$$\left(\begin{array}{cc|c} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array}\right)$$

Note how the augumented matrix only contains the coefficients of each term – the equal sign is replaced with a vertical line. Similarly, the coefficient matrix,

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

contains all the coefficients of the system of equations.
The solution of such a system is commonly expressed as a vector:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

Gauss-Jordan Elimination

Definition — A matrix is known to be in *row-echelon form* if

1. All rows consisting of only zeroes are at the bottom
2. The leading entry (that is the left-most nonzero entry) of every nonzero row is to the right of the leading entry of every row above

An example of a matrix in row-echelon form is shown below:

$$\begin{pmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 2 & a_4 & a_5 \\ 0 & 0 & 0 & 1 & a_6 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

However, when solving systems, we are more concerned with the *reduced* row-echelon form.

Reduced Row-Echelon form — A matrix is said to be in rref (*reduced row-echelon form*) when it satisfies all the following conditions:

1. It is in row echelon form.
2. The leading entry in each nonzero row is a 1 (called a leading 1).
3. Each column containing a leading 1 has zeros in all its other entries.

An example of such is shown below (note that it is not always an identity matrix):

$$\begin{pmatrix} 1 & 0 & a_1 & 0 & b_1 \\ 0 & 1 & a_2 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{pmatrix}$$

Gauss-Jordan elimination, also known as gaussian elimination or row reduction, is the method in which we transform a matrix into rref (reduced row-echelon) form.

It uses only the elementary row operations, shown below:

- Divide a row by a non-zero scalar
- Subtract a multiple of a row from another row
- Swap two rows

Example — Write the augumented matrix and rref for the following matrix:

$$\begin{cases} x + 3y + z = 9 \\ x + y - z = 1 \\ 3x + 11y + 5z = 35 \end{cases}$$

What are the solutions for x, y, and z?

Solution.

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 1 & 1 & -1 & 1 \\ 3 & 11 & 5 & 35 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 2 & 2 & 8 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -2 & -3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

When put in rref form, it is clear to see that this system

$$\begin{cases} x - 2z = -3 \\ y + z = 4 \end{cases}$$

has infinite solutions. For a certain real number t, we could write the solutions as such:

$$\boxed{\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2t - 3 \\ -t + 4 \\ t \end{pmatrix}}$$

□

On the Solutions of Linear Systems; Matrix Algebra

Definition — The *rank* of a matrix A is the number of leading 1's in $\text{rref}(A)$, denoted by $\text{rank}(A)$.

Number of Solutions A system of equations is *consistent* if it has at least one solution; else it is *inconsistent*. Note the following truths: For a system of n linear equations with m variables, with a coefficient matrix of size $n \times m$,

(a) $\text{rank}(A) \leq n, m$

(b) if $\text{rank}(A) = n$, the system is consistent

(c) if $\text{rank}(A) < m$, then the system has no sol or inf sols

(d) if $\text{rank}(A) = m$, then the system has no sol or one sol

Note that if a linear system has exactly one solution, then there must be at least as many equations as there are variables. This satisfies the truths above; $m = \text{rank}(A) \leq n \implies m \leq n$.

Theorem

A linear system of n equations in n variables has a unique solution if and only if the rank of its coefficient matrix A is n. In this case,

$$\text{rref}(A) = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Matrix Algebra

Sum of Matrices — The sum of 2 matrices of the same size is given by

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{pmatrix}$$

Scalar Multiples of Matrices — The product of a scalar with a matrix is given by

$$k\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} ka_{11} & \cdots & ka_{1m} \\ \vdots & \ddots & \vdots \\ ka_{n1} & \cdots & ka_{nm} \end{pmatrix}$$

Matrix-Vector Multiplicaton — This can be seen in two different ways:
<u>Row-wise</u>: Given a matrix $A \in \mathbb{R}^{n\times m}$ with row vectors $\vec{\mathbf{w}}_1, \ldots, \vec{\mathbf{w}}_n$ and vector $\vec{\mathbf{x}} \in \mathbb{R}^m$, then

$$A\vec{\mathbf{x}} = \begin{pmatrix} - & \vec{\mathbf{w}}_1 & - \\ & \vdots & \\ - & \vec{\mathbf{w}}_n & - \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \vec{\mathbf{w}}_1 \cdot \vec{\mathbf{x}} \\ \vdots \\ \vec{\mathbf{w}}_n \cdot \vec{\mathbf{x}} \end{pmatrix}$$

<u>Column-wise</u>: Given a matrix $A \in \mathbb{R}^{n\times m}$ with column vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m$ and vector $\vec{\mathbf{x}} \in \mathbb{R}^m$, then

$$A\vec{\mathbf{x}} = \begin{pmatrix} | & & | \\ \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_m \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = x_1\vec{\mathbf{v}}_1 + \cdots + x_m\vec{\mathbf{v}}_m$$

Definition — A vector $\vec{\mathbf{b}} \in \mathbb{R}^n$ is called a <u>linear combination</u> of the vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m$ in $\mathbb{R}^n$ if there exist scalars $x_1, \ldots, x_m$ such that

$$\vec{\mathbf{b}} = x_1\vec{\mathbf{v}}_1 + \cdots + x_m\vec{\mathbf{v_m}}$$

Theorem (Algebraic Rules)

If A is an $n \times m$ matrix, $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ are vectors in $\mathbb{R}^m$, and k is a scalar, then

a) $A(\vec{\mathbf{x}} + \vec{\mathbf{y}}) = A\vec{\mathbf{x}} + A\vec{\mathbf{y}}$

b) $A(k\vec{\mathbf{x}}) = k(A\vec{\mathbf{x}})$

Note that if given a systme defined by the augumented matrix

$$\left(A \;\middle|\; \vec{\mathbf{b}}\right)$$

we can rewrite it to an equivalent form,

$$A\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

which, solving for $\vec{\mathbf{x}}$ results in

$$\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$$

The invertibility of matrices will be discussed in the next section.

§2 Linear Transformations

Introduction to Linear Transformations and Their Inverses

Definition — A function $T : \mathbb{R}^m \to \mathbb{R}^n$ is a *linear transformation* if there exists an $n \times m$ matrix A such that

$$T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}}$$

for all $\vec{\mathbf{x}}$ in the vector space $\mathbb{R}^m$.

A linear transformation can be thought of as a function; if we write $T(\vec{\mathbf{x}}) = \vec{\mathbf{y}}$, then we get

$$\vec{\mathbf{y}} = A\vec{\mathbf{x}}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n_1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} a_{11}x_1 & \cdots & a_{1m}x_m \\ \vdots & \ddots & \vdots \\ a_{n1}x_1 & \cdots & a_{nm}x_m \end{pmatrix}$$

This more visually shows the transformation from the point $(x_1, \ldots, x_m) \in \mathbb{R}^m$ to the point $(y_1, \ldots, y_n) \in \mathbb{R}^n$ with linear systems

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1m}x_m \\ \vdots &= \ \vdots \ \ + \ \vdots \ + \ \ \vdots \\ y_n &= a_{n1}x_1 + \cdots + a_{nm}x_m \end{aligned}$$

Theorem (Linear Transformations)

A transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is linear if and only if

a) $T(\vec{\mathbf{v}} + \vec{\mathbf{w}}) + T(\vec{\mathbf{v}}) + T(\vec{\mathbf{w}}) \quad \forall \quad \vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathbb{R}^m$

b) $T(k\vec{\mathbf{v}}) = kT(\vec{\mathbf{v}}) \quad \forall \quad \vec{\mathbf{v}} \in \mathbb{R}^m$ and $k \in \mathbb{R}$

note: distribution vectors and transition matrices $\longrightarrow$ basically markov chains

Linear Transformations in Geometry

There are five main different types of transformations that 2×2 matrices can perform.

Scalings — For any positive constant k, the matrix

$$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

defines scaling by k.

Projections — The *orthogonal projection* of $\vec{\mathbf{x}} \in \mathbb{R}^2$ onto a line L in the coordinate plane, is

$$\vec{\mathbf{x}}^{\parallel} = \text{proj}_L(\vec{\mathbf{x}}) = \left(\frac{\vec{\mathbf{x}} \cdot \vec{\mathbf{w}}}{\vec{\mathbf{w}} \cdot \vec{\mathbf{w}}}\vec{\mathbf{w}}\right) = (\vec{\mathbf{x}} \cdot \vec{\mathbf{u}})\vec{\mathbf{u}}$$

if $\vec{\mathbf{w}}$ is a nonzero vector parallel to L and $\vec{\mathbf{u}}$ is the unit vector parallel to L.
The transformation $T(\vec{\mathbf{x}}) = \text{proj}_L(\vec{\mathbf{x}})$ is linear, with matrix

$$P = \frac{1}{w_1^2 + w_2^2}\begin{pmatrix} w_1^2 & w_1w_2 \\ w_1w_2 & w_2^2 \end{pmatrix} = \begin{pmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{pmatrix}$$

Reflections — Given a line L in the coordinate plane, and a unit vector $\vec{\mathbf{u}}$ parallel to L, the *reflection of* $\vec{\mathbf{x}}$ *about* L is given by

$$\text{ref}_L(\vec{\mathbf{x}}) = 2\text{proj}_L(\vec{\mathbf{x}}) - \vec{\mathbf{x}} = 2(\vec{\mathbf{x}} \cdot \vec{\mathbf{u}})\vec{\mathbf{u}} - \vec{\mathbf{x}} = \begin{pmatrix} 2u_1^2 - 1 & 2u_1u_2 \\ 2u_1u_2 & 2u_2^2 - 1 \end{pmatrix}$$

The matrix of the transformation is in the form

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

where $a^2 + b^2 = 1$.

Rotations — The matrix of a counterclockwise rotation in $\mathbb{R}^2$ through an angle θ is

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

The matrix of this transformation is in teh form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

where $a^2 + b^2 = 1$.

Shears — There are two types of shears: horizontal and vertical. Given an arbitrary constant k, the matrix of a horizontal shear is in the form

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

and the matrix of a vertical shear is of the form

$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$$

Matrix Products

Definition — Let B be an $n \times p$ matrix and A a $q \times m$ matrix. The product BA is defined if and only if $p = q$. The product BA is defined as the matrix of the linear transformation $T(\vec{\mathbf{x}}) = B(A\vec{\mathbf{x}})$. We can write that

$$T(\vec{\mathbf{x}}) = B(A\vec{\mathbf{x}}) = (BA)\vec{\mathbf{x}}$$

for all $\vec{\mathbf{x}} \in \mathbb{R}^m$, which means that the product BA is an $n \times m$ matrix.

Matrix multiplication has different properties than normal multiplication; namely, matrix multiplication is

1. **Non-Commutative**
   $AB \neq BA$ in general, but when $AB = BA$ then the two matrices A and B *commute*.
2. **Associative**
   $(AB)C = A(BC) = ABC$. Matrix multiplication is associative.
3. **Distributive**
   If A and B are $n \times p$ matrices, and C and D are $p \times m$ matrices, then
   $$A(C + D) = AC + AD \qquad \text{and} \qquad (A + B)C = AC + BC$$
   If A is an $n \times p$ matrix, B is a $p \times m$ matrix, and k is a scalar, then
   $$(kA)B = A(kB) = k(AB)$$

Theorem (Multiplying with an Identity Matrix)
For an $n \times m$ matrix A
$$AI_m = I_nA = A$$

The Inverse of a Linear Transformation

Definition — A function T from X to Y is called invertible if the equation $T(x) = y$ has a unique solution x in X for each y in Y. This means that for all $x \in X$, $T^{-1}(T(x)) = x$

A square matrix A is said to be *invertible* if the linear transformation

$$\vec{\mathbf{y}} = T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}}$$

is invertible, with an inverse of

$$\vec{\mathbf{x}} = T^{-1}(\vec{\mathbf{y}}) = A^{-1}\vec{\mathbf{y}}$$

Theorem (Invertibility)
An $n \times n$ matrix A is invertible if and only if

$$\text{rref}(A) = I_n \qquad \text{or (equivalently)} \qquad \text{rank}(A) = n$$

Note that this makes solving linear systems straightforward, as for a system

$$A\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

the solution would be

$$\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$$

Example — Is the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{pmatrix}$$

invertible? If so, find its' inverse.

Solution. We first confirm that the matrix is invertible; solving for the rref gives us I_3, which means it is invertible.
To find the inverse, we find the inverse of the linear transformation

$$\vec{\mathbf{y}} = A\vec{\mathbf{x}} \quad \Longrightarrow \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + x_3 \\ 2x_1 + 3x_2 + 2x_3 \\ 3x_1 + 8x_2 + 2x_3 \end{pmatrix}$$

Solving for x_1, x_2, and x_3 gives us

$$\begin{aligned} x_1 &= 10y_1 - 6y_2 + y_3 \\ x_2 &= -2y_1 + y_2 \\ x_3 &= -7y_1 + 5y_2 - y_3 \end{aligned}$$

which means that

$$\boxed{A^{-1} = \begin{pmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{pmatrix}}$$

Putting this in matrix form, we essentially shifted from

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 2 & 8 & 2 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array}\right)$$

with the elementary row operations. □

This ties into how to find the inverse of a matrix;

Finding the Inverse — To find the *inverse* of an $n \times n$ matrix A, form the $n \times (2n)$ matrix $\left(A \mid I_n\right)$ and compute hte rref of that matrix.
If $\text{rref}\left(A \mid I_n\right)$ is of the form $\left(I_n \mid A\right)$, then A is invertible, and $A^{-1} = B$. Otherwise, if its of anohter form, then A is not invertible.

Theorem (Multiplying with Inverse)
For an invertibel $n \times n$ matrix A,

$$A^{-1}A = AA^{-1} = I_n$$

Theorem (Inverse of Product)
If A and B are invertible $n \times n$ matrices, hten BA is invertible as well, and

$$(BA)^{-1} = A^{-1}B^{-1}$$

Inverse and Determinant of 2 x 2 Matrix — The 2×2 matrix has an inverse

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

only when $\det(A) \neq 0$, where $\det(A)$ is the *determinant* of matrix A, equal to

$$\det(A) = ad - bc$$

Geometrical Interpretation of Determinant of a 2 x 2 Matrix

If

$$A = \begin{pmatrix} \vec{\mathbf{v}} & \vec{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a 2×2 matrix with nonzero columns $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$, then

$$\det A = \|\vec{\mathbf{v}}\| \sin(\theta) \|\vec{\mathbf{w}}\| = ad - bc$$

where θ is the oriented angle from $\vec{\mathbf{v}}$ to $\vec{\mathbf{w}}$. We can see that the determinant corresponds with the area of the parallelogram spanned by the two vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$.

§3 Subspaces of $\mathbb{R}^n$ and Their Dimensions

Image and Kernal of a Linear Transformation

Definition — The *image* of a function contains all the values the function takes in its target space. If f is a function from X to Y, then

$$\text{image}(f) = \{f(x) : x \text{ in } X\}$$

See 1.

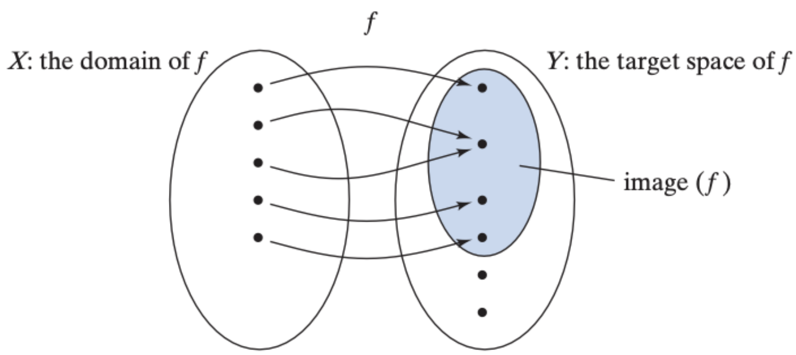

Figure 1: image of a function visualized

Sometimes, people will refer to the image of a function as its *range*, but other times, they will refer to the target space as *range* – however, this is will not be used with this textbook.

Definition — Given vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_m \in \mathbb{R}^n$, the set of all linear combinations $c_1\vec{\mathbf{v}}_1 + \cdots + c_m\vec{\mathbf{v}}_m$ of those vectors is called their *span*.

$$\text{span}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_m) = \{c_1\vec{\mathbf{v}}_1 + \cdots + c_m\vec{\mathbf{v}}_m : c_1, \dots, c_m \in \mathbb{R}\}$$

Theorem

The image of a linear transformaiton $T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}}$ is the span of the column vectors of A. We denote the image of T by $\text{im}(T)$ or $\text{im}(A)$.

Some properties of an image:

- The zero vector $\vec{\mathbf{0}}$ in $\mathbb{R}^n$ is in the image of T
- If $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ are in the image of T, then so is $\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2$
- If $\vec{\mathbf{v}}$ is in the image of T and k is an arbitrary scalar, then $k\vec{\mathbf{v}}$ is in the image of T as well

Definition — The *kernel* of a linear transformation $T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}}$ from $\mathbb{R}^m$ to $\mathbb{R}^n$ consists of all zeros of the transformation, or in other words, the solution set to the linear system

$$A\vec{\mathbf{x}} = 0$$

We denote the kernel of T by $\ker(T)$ or $\ker(A)$.

Some properties of a kernel:

- The zero vector $\vec{\mathbf{0}}$ in $\mathbb{R}^m$ is in the kernel of T
- The kernel is closed under addition
- The kernel is closed under scalar multiplication

Notice how for a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$,

- $\text{im}(T) = \{T(\vec{\mathbf{x}}) : \vec{\mathbf{x}} \in \mathbb{R}^m\}$ is a subset of the *target space* $\mathbb{R}^n$ of T
- $\ker(T) = \{\vec{\mathbf{x}} \in \mathbb{R}^m : T(\vec{\mathbf{x}}) = 0\}$ is a subset of the *domain* $\mathbb{R}^m$ of T

Subspaces of $\mathbb{R}^n$; Bases and Linear Independence

Definition — A subset W of the vector space $\mathbb{R}^n$ is called a (linear) *subspace* of $\mathbb{R}^n$ if it has the following three properties:

a) W contains the zero vector in $\mathbb{R}^n$

b) W is closed under addition: If $\vec{\mathbf{w}}_1$ and $\vec{\mathbf{w}}_2$ are both in W, then so is $\vec{\mathbf{w}}_1 + \vec{\mathbf{w}}_2$

c) W is closed under scalar multiplication: If $\vec{\mathbf{w}}$ is in W and k is an arbitrary scalar, the $k\vec{\mathbf{w}}$ is in W

This means that if $T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}}$ is a linear transformation from $\mathbb{R}^m \to \mathbb{R}^n$, then

- $\ker(T) = \ker(A)$ is a subspace of $\mathbb{R}^m$
- $\text{image}(T) = \text{im}(A)$ is a subspace of $\mathbb{R}^n$

Definition — Consider vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m \in \mathbb{R}^n$

a. We say that a vector $\vec{\mathbf{v}}_i$ in the list $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m$ is *redundant* if $\vec{\mathbf{v}}_i$ is a linear combination of the preceding vectors

b. The vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m$ are called *linearly independent* if none of the are redundant; else, they're *linearaly dependent*

c. We say that the vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m$ in a subspace V of $\mathbb{R}^n$ form a *basis* of V if they span V and are linearly independent

When constructing the basis of images, we must omit the redundant vectors.

Example — Are the following vectors linearly independent?

$$\vec{\mathbf{v}}_1 = \begin{pmatrix}1\\2\\3\end{pmatrix} \qquad \vec{\mathbf{v}}_2 = \begin{pmatrix}4\\5\\6\end{pmatrix} \qquad \vec{\mathbf{v}}_3 = \begin{pmatrix}7\\8\\9\end{pmatrix}$$

Solution. We have to see whether there exists a c_1 and c_2 such that $\vec{\mathbf{v}}_3 = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2$. With the augmented matrix

$$\left(\begin{array}{cc|c}1 & 4 & 7\\2 & 5 & 8\\3 & 6 & 9\end{array}\right)$$

we find the unique colution $c_1 = -1$, $c_2 = 2$, which makes these vectors linearly dependent. □

In a *linear relation*

$$c_1\vec{\mathbf{v}}_1 + \cdots + c_m\vec{\mathbf{v}}_m = \vec{\mathbf{0}}$$

the *trivial relation* is known as the case where $c_1 = \cdots = c_m = 0$.

Theorem
The column vectors of an $n \times m$ matrix A are linearly independent if and only if $\ker(A) = \{\vec{\mathbf{0}}\}$, or equivalently if $\text{rank}(A) = m$.

The Dimension of a Subspace of $\mathbb{R}^n$

Theorem
All bases of a subspace V of $\mathbb{R}^n$ consist of the same number of vectors.

The basis of a line is just one vector; the basis of a plane is two vectors, and the basis of $\mathbb{R}^3$ is jsut 3 vectors.

Definition — Consider a subspace V of $\mathbb{R}^n$. The number of vectors in a basis of V is called teh *dimension* of V, denoted by $\dim(V)$.

Basically, the basis of a subspace is the set of linearly independent vectors that span the subspace. The number of these vectors form the dimension.

Example — Given the matrix

$$A = \begin{pmatrix} 1 & 2 & 2 & -5 & 6 \\ -1 & -2 & -1 & 1 & -1 \\ 4 & 8 & 5 & -8 & 9 \\ 3 & 6 & 1 & 5 & -7 \end{pmatrix}$$

Find the basis of the kernel of A, and the dimension of the kernel, and find the basis of the image of A, and the dimension of the image.

Solution. For the first part, to find the kernel, we'll solev the linear system $A\vec{\mathbf{x}} = 0$. We get that

$$B = \text{rref}(A) = \begin{pmatrix} 1 & 2 & 0 & 3 & -4 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Writing out, we can see that the solutions are in the form

$$\vec{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2s - 3t + 4r \\ s \\ 4t - 5r \\ t \\ r \end{pmatrix} = s\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix} + r\begin{pmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{pmatrix}$$

We can confirm that these 3 vectors which span the subspace are linearly independent, meaning they form the basis for the subspace, which has 3 dimensions.
For the second part, the image of A, we just have to take the column vectors and weed out the redundant ones. Note how we can form a relation between the redundant columns of A and the ones that of B, the rref of A.

From there, we find that only the first and third column of B, which means only the first and third columns of A, form the basis fo A, so the dimension is 2 and the basis are

$$\begin{pmatrix} 1 \\ -1 \\ 4 \\ 3 \end{pmatrix} \quad \begin{pmatrix} 2 \\ -1 \\ 5 \\ 1 \end{pmatrix}$$

□

From this example, we can see the following thm:

Theorem

To construct a basis for the image of A, pick the column vectors of A the correspond to teh columns of $\text{rref}(A)$ containing the leading 1's.

This leads to the following realization: for any $n \times m$ matrix A,

$$\dim(\text{im}\, A) = \text{rank}(A)$$

We can also observe that

$$\dim(\ker A) = m - \text{rank}(A)$$

(note this in the prev example). Note that by adding the two values up,

$$\dim(\text{im}\, A) + \dim(\text{im}\, A) = m$$

The dimension of the kernel of A is known as the *nullity* fo A; which leads us to teh important theorem, the

Theorem (Rank-Nullity Theorem)

For any $n \times m$ matrix A,

$$(\text{nullity of } A) + (\text{rank of } A) = m$$

Note that the kernel can be found if there is lienarly dependent redundant vectotrs, and writing the linear combinatoni that leads to the redundancy can be used to find the 0, helping find the kernel.

Coordinates

Coordinates in Subspace — Consider a basis $\mathcal{B} = (\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \dots, \vec{\mathbf{v}}_m)$ of a subspace $V \in \mathbb{R}^n$. Given a vector $\vec{\mathbf{x}} \in V$, it can be written uniquely as

$$\vec{\mathbf{x}} = c_1\vec{\mathbf{v_1}} + c_2\vec{\mathbf{v}}_2 + \cdots + c_m\vec{\mathbf{v}}_m$$

The scalars $c_1, c_2, \dots, c_m$ are called the *$\mathcal{B}$-coordinates* of $\vec{\mathbf{x}}$, and the vector

$$[\vec{\mathbf{x}}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

is the *$\mathcal{B}$-coordinate vector* of $\vec{\mathbf{x}}$.

Note that we can write that

$$\vec{\mathbf{x}} = c_1\vec{\mathbf{v}}_1 + \cdots + c_m\vec{\mathbf{v}}_m = S\,[\vec{\mathbf{x}}]_{\mathcal{B}} \qquad \text{where} \quad S = \begin{pmatrix} | & & | \\ \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_m \\ | & & | \end{pmatrix}$$

Theorem (Linearity of Coordinates)

If $\mathcal{B}$ is a basis of a subspace V of $\mathbb{R}^n$, then for all vectors $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in V$ and $k \in \mathbb{R}$,

1. $[\vec{\mathbf{x}} + \vec{\mathbf{y}}]_{\mathcal{B}} = [\vec{\mathbf{x}}]_{\mathcal{B}} + [\vec{\mathbf{y}}]_{\mathcal{B}}$
2. $[k\vec{\mathbf{x}}]_{\mathcal{B}} = k\,[\vec{\mathbf{x}}]_{\mathcal{B}}$

Transformations in Coordinates — Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ in a basis $\mathcal{B} = (\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_n)$ of $\mathbb{R}^n$. We can write a unique $n \times n$ matrix B, called the *$\mathcal{B}-$matrix of T*, such that

$$[T(\vec{\mathbf{x}})]_{\mathcal{B}} = B\,[\vec{\mathbf{x}}]_{\mathcal{B}}$$

for all $\vec{\mathbf{x}} \in \mathbb{R}^n$.

Note that we can construct B as

$$B = \begin{pmatrix} | & & | \\ [T(\vec{\mathbf{v}}_1)]_{\mathcal{B}} & \cdots & [T(\vec{\mathbf{v}}_n)]_{\mathcal{B}} \\ | & & | \end{pmatrix}$$

Theorem

Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ and a basis $\mathcal{B} = (\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_n)$, let A be the standard matrix of T, B be the $\mathcal{B}$-matrix of T, and S be the basis of $\mathcal{B}$ as the column vectors of the matrix

$$S = \begin{pmatrix} | & & | \\ \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_n \\ | & & | \end{pmatrix}$$

Then, we have that

$$AS = SB$$

and following,

$$B = S^{-1}AS \quad \text{and} \quad A = S^{-1}BS$$

Definition — Given two $n \times n$ matrices A and B, they're similar if there exists an invertible S such that

$$AS = SB$$

Sometiems, to make a problem simpler, we can shift basis to make linear transformations simpler (similar to swithcing coordinate planes in some physics problems). Essentially, we would like for a transformation T to be such that the matrix B that represents this transformation in a basis $\mathcal{B}$ is in the form

$$B = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}$$

This is simply a scaling matrix – except scaling the axis of the revised basis. To formalize, for a certain basis $\mathcal{B} = (\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n)$, the $\mathcal{B}$-matrix of T is diagonal if and only if

$$T(\vec{\mathbf{v}}_i) = c_i \vec{\mathbf{v}}_i$$

for scalars $c_1, \dots, c_n$.
This relates to the topic of *eigenvalues* and *eigenbasis*, which will be touched upon later.

§4 Orthogonality and Least Squares

Orthogonal Projects and Orthonormal Bases

Definition — The vectors $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_m \in \mathbb{R}^n$ are called *orthonormal* if tehy are all unit vectors and orthogonal to each other.

Notice that orthonormal vectors

1. are linearly independent

2. form the basis of the subspace

Orthogonal Projection — Consider a vector $\vec{\mathbf{x}} \in \mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. We can write

$$\vec{\mathbf{x}} = \vec{\mathbf{x}}^{\|} + \vec{\mathbf{x}}^{\perp}$$

The vector $\vec{\mathbf{x}}^{\|}$ is called the *orthogonal projection* of $\vec{\mathbf{x}}$ onto V, denoted by $\text{proj}_v \vec{\mathbf{x}}$. This transformation is linear.
If V has orthonormal basis $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_m$, then

$$\text{proj}_V \vec{\mathbf{x}} = \vec{\mathbf{x}}^{\|} = (\vec{\mathbf{u}}_1 \cdot \vec{\mathbf{x}})\vec{\mathbf{u}}_1 + \cdots + (\vec{\mathbf{u}}_m \cdot \vec{\mathbf{x}})\vec{\mathbf{u}}_m$$

for all $\vec{\mathbf{x}} \in \mathbb{R}^n$.

blah blah bla

Gram-Schmidt Process and QR Factorization

Gram-Schmidt Process — Consider a basis $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_m$ of a subspace V of $\mathbb{R}^n$. We can write, for a j in $2, \dots, m$,

$$\vec{\mathbf{v}}_j^{\perp} = \vec{\mathbf{v}}_j - \vec{\mathbf{v}}_j^{\|} = \vec{\mathbf{v}}_j - (\vec{\mathbf{v}}_j \cdot \vec{\mathbf{u}}_1)\vec{\mathbf{u}}_1 - \cdots - (\vec{\mathbf{v}}_j \cdot \vec{\mathbf{u}}_{j-1})\vec{\mathbf{u}}_{j-1}$$

where

$$\vec{\mathbf{u}}_i = \frac{\vec{\mathbf{v}}_i}{\|\vec{\mathbf{v}}_i\|}$$

Geometrically, this is just subtracting the vector from all orthonormal projections of itself on lower dimension vectors, such that we'll end up with a resulting vector orthogonal to all previous vectors.

QR Factorization — Note that we can represent the change in basis from $\mathcal{B} = (\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m)$ to the basis $\mathcal{U} = (\vec{\mathbf{u}}_1, \ldots, \vec{\mathbf{u}}_m)$, with

$$\underbrace{\begin{pmatrix} | & & | \\ \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_m \\ | & & | \end{pmatrix}}_{M} = \underbrace{\begin{pmatrix} | & & | \\ \vec{\mathbf{u}}_1 & \cdots & \vec{\mathbf{u}}_m \\ | & & | \end{pmatrix}}_{Q} R$$

where R is the change of basis matrix.
The entries of R are given as

1. $r_{11} = \|\vec{\mathbf{v}}_1\|$
2. $r_{ij} = \vec{\mathbf{u}}_i \cdot \vec{\mathbf{v}}_j$ for $i < j$
3. $r_{jj} = \left\|\vec{\mathbf{v}}_j^{\perp}\right\|$
4. for $i > j$, $r_{ij} = 0$

An example of what the matrix R would look like is as follows:

$$R = \begin{pmatrix} \|\vec{\mathbf{v}}_1\| & \vec{\mathbf{u}}_1 \cdot \vec{\mathbf{v}}_2 & \vec{\mathbf{u}}_1 \cdot \vec{\mathbf{v}}_3 \\ 0 & \left\|\vec{\mathbf{v}}_2^{\perp}\right\| & \vec{\mathbf{u}}_2 \cdot \vec{\mathbf{v}}_3 \\ 0 & 0 & \left\|\vec{\mathbf{v}}_3^{\perp}\right\| \end{pmatrix}$$

Note how by multiplying the matrices QR, we get a something similar to

$$\begin{aligned} &(\vec{\mathbf{u}}_1 \cdot \vec{\mathbf{v}}_j)\vec{\mathbf{u}}_1 + \cdots + (\vec{\mathbf{u}}_{j-1} \cdot \vec{\mathbf{v}}_j)\vec{\mathbf{u}}_{j-1} + \|\vec{\mathbf{v}}_j\|\vec{\mathbf{u}}_j \\ &= \vec{\mathbf{v}}_j^{\parallel} + \vec{\mathbf{v}}_j^{\perp} = \vec{\mathbf{v}}_j \end{aligned}$$

Orthogonal Transformations and Orthogonal Matrices

Definition — A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is called *orthogonal* if

$$\|T(\vec{\mathbf{x}})\| = \|\vec{\mathbf{x}}\|$$

If $T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}}$ is an orthogonal transformation, then we can also say A is an *orthogonal matrix*.

Theorem

Given an orthogonal transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, if vectors $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathbb{R}^n$ are orthogonal, then so are $T(\vec{\mathbf{v}})$ and $T(\vec{\mathbf{w}})$.

The rotation matrix is an example of an orthogonal transformation.
Note that following this, we can see that

1. a linear transformation T is orthogonal if and only if $T(\vec{\mathbf{e}}_1), \ldots, T(\vec{\mathbf{e}}_n)$ form an orthonormal basis in $\mathbb{R}^n$
2. an $n \times n$ matrix A is orthogonal if and only if its columns form an orthonormal basis in $\mathbb{R}^n$

Note that, following this, we can also show that

1. The product AB of two orthogonal $n \times n$ matrices A and B is orthogonal
2. the inverse A^{-1} of an orthogonal $n \times n$ matrix A is orthogonal

Definition — Given an $m \times n$ matrix A, the *transpose* A^T of A is the $n \times m$ matrix who's ijth entry is the jith entry of A.
If A is a square matrix, it is *symmetric* if $A^T = A$, and *skew-symmetric* if $A^T = -A$.

Note that for a column vector

$$\vec{\mathbf{v}} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \text{ then } \vec{\mathbf{v}}^T = \begin{pmatrix} a & b & c \end{pmatrix}$$

This means that for two column vectors $\vec{\mathbf{v}}, \vec{\mathbf{w}}$, we can convert the dot product into matrix multiplication

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \vec{\mathbf{v}}^T \vec{\mathbf{w}}$$

Theorem

An $n \times n$ matrix A is orthogonal if and only if

$$A^T A = I_n$$

The transpose has hte following properties:

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

b) $(kA)^T = kA^T$

c) $(AB)^T = B^T A^T$

d) rank(A^T)=rank(A)

e) $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Note that since we can write

$$\begin{aligned} \text{proj}_V \vec{\mathbf{x}} &= \vec{\mathbf{u}}_1(\vec{\mathbf{u}}_1 \cdot \vec{\mathbf{x}}) + \cdots + \vec{\mathbf{u}}_m(\vec{\mathbf{u}}_m \cdot \vec{\mathbf{x}}) \\ &= (\vec{\mathbf{u}}_1 \vec{\mathbf{u}}_1^T + \cdots + \vec{\mathbf{u}}_m \vec{\mathbf{u}}_m^T)\vec{\mathbf{x}} \end{aligned}$$

leading to the following,

Matrix of an Orthogonal Projection — Given a subspace $V \in \mathbb{R}^n$ with an orthonormal basis $\vec{\mathbf{u}}_1, \ldots, \vec{\mathbf{u}}_m$, the matrix P of the projection onto V is given by

$$P = QQ^T \qquad \text{where} \qquad Q = \begin{pmatrix} | & & | \\ \vec{\mathbf{u}}_1 & \cdots & \vec{\mathbf{u}}_m \\ | & & | \end{pmatrix}$$

Note that P is a square matrix, and symmetric.

Least Squares and Data Fitting

Theorem

For any matrix A,

$$(\text{im} A)^\perp = \ker\left(A^T\right)$$

This makes sense, as the space perpendicular to the image means the dot product is 0, which means the transpose is 0 which means its the kernel.

Theorem

If A is an $n \times m$ matrix,

$$\ker(A) = \ker(A^T A)$$

Furthermore, if $\ker(A) = \{\vec{\mathbf{0}}\}$, then $A^T A$ is invertible.

Definition — Consider a linear system

$$A\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

where A is an $n \times m$ matrix. A vector $\vec{\mathbf{x}}^*$ is called a *least-squares solution* of this system if $\left\|\vec{\mathbf{b}} - A\vec{\mathbf{x}}^*\right\| \leq \left\|\vec{\mathbf{b}} - A\vec{\mathbf{x}}\right\| \quad \forall \quad \vec{\mathbf{x}} \in \mathbb{R}^n$

Basically, the least squares solution solves the inconsistent system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$, for the closest solution for $\vec{\mathbf{x}}$ to which $A\vec{\mathbf{x}}$ gets to $\vec{\mathbf{b}}$.

Theorem (The Normal Equation)

The least-squares solutions of the system

$$A\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

are the exact solutions of the consistent system

$$A^T A\vec{\mathbf{x}} = A^T\vec{\mathbf{b}}$$

which is called the *normal equation* fo the original system.

Note that if $\ker(A) = \{\vec{\mathbf{0}}\}$, then the matrix $A^T A$ is invertible, and this linear system has the unique solution

$$\vec{\mathbf{x}}^* = \left(A^T A\right)^{-1} A^T\vec{\mathbf{b}}$$

Note the similarity finding the least squares equation has with the projection – the projection of a vector $\vec{\mathbf{b}}$ onto a plane V is essentially finding the shortest possible length of $\vec{\mathbf{x}} - \vec{\mathbf{v}}$ for any vector $\vec{\mathbf{v}}$ on the plane V, for then that vector $\vec{\mathbf{v}}$ will be the projection.

Matrix of Orthogonal Projection — Consider a subspace V of $\mathbb{R}^n$ with basis $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_m$. Let

$$A = \begin{pmatrix} | & & | \\ \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_m \\ | & & | \end{pmatrix}$$

Then the matrix of the orthogonal projection onto V is

$$A\left(A^T A\right)^{-1} A^T$$

Note that if the basis vectors are orthonormal, then $A^T A = I_m$, which'll simplify the above equation to having the orthogonal projection just be AA^T.

§5 Determinants

Introduction to Determinants

Patterns, Inversions, Determinant — A *pattern* in an $n \times n$ matrix A is a way to choose n entries of the matrix so that there is one chosen entry in each row and in each column of A. Two entries in a pattern are said to be *inverted* when one of them is to the right and above the other.
The determinant of A is then defined as the sum of all pattern products with an even number of inversions, subtracted from all the patterns products with an odd number of inversions, with the following formula:

$$\det A = \sum (-1)^{\#\text{ of inversions}} (\text{prod } P)$$

Properties of Determinants

Theorem
If A is a square matrix, then

$$\det\left(A^T\right) = \det A$$

Cramer's Rule — For a linear system

$$A\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

where A is an invertible $n \times n$ matrix, the solution vector is given by

$$x_i = \frac{\det\left(A_{\vec{\mathbf{b}},i}\right)}{\det A}$$

where $A_{\vec{\mathbf{b}},i}$ is the matrix after replacing the ith column of A by $\vec{\mathbf{b}}$.

Theorem
For an invertible $n \times n$ matrix A, the *classical adjoint* $\operatorname{adj}(A)$ is the $n \times n$ matrix which is the tranpose of the cofactor array, $(-1)^{i+j}\det\left(A_{ij}\right)$. With this,

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

§6 Eigenvalues and Eigenvectors

Diagonalization

§7 Appendix

note: remember to cite wikipedia and otto linear algebra fundamentals 5th ed