Mathematical Methods Notes

Hashan Punchihewa

https://hashanp.xyz

Contents

Chapter 1

Mathematical Methods

1.1 Sequences

Proposition 1.1.1

1. $|x| < a \leftrightarrow -a < x < a$
2. $|xy| = |x||y|$
3. $|x/y| = |x|/|y|$
4. $|x+y| \leq |x| + |y|$ *(Triangle Inequality)*
5. $|x-y| \geq ||x| - |y||$

Definition 1.1.1

If S is a set of real numbers:

1. *u is an upper bound on S if for any $u \geq s$, for all $s \in S$.*
2. *l is a lower bound on S if for any $l \leq s$, for all $s \in S$.*
3. *S is bounded above if it has an upper bound.*
4. *S is bounded below if it has a lower bound.*
5. *S is bounded if it has a lower and upper bound.*
6. *$\sup(S)$, called the supremum of S is an upper bound of S such that there is no smaller upper bound of S. If there is no upper bound, we say $\sup(S) = \infty$.*
7. *$\inf(S)$, called the infimum of S is a lower bound of S such that there is no larger lower bound of S. If there is no lower bound, we say $\inf(S) = -\infty$.*

Definition 1.1.2 *A sequence a_n, for $n \geq 1$ converges to a limit $l \in \mathbb{R}$, denoted as:*

$$\lim_{n\to\infty} a_n = l$$

Or $a_n \to l$ as $n \to \infty$, if for all $\epsilon > 0$, there exists a positive integer N, such that for all $n > N$, $|a_n - l| < \epsilon$.

Definition 1.1.3 *A sequence diverges, if there is no limit to which it converges.*

Proposition 1.1.2

1. *As $n \to \infty$, $1/n^k \to 0$, where $k \neq 0$.*
2. *As $n \to \infty$, $1/\ln n \to 0$, where $n > 1$.*

Proof:

1.

$$1/n^k < \epsilon \tag{1.1}$$
$$1 < n^k \epsilon \tag{1.2}$$
$$1/\epsilon < n^k \tag{1.3}$$
$$1/\sqrt[k]{\epsilon} < n \tag{1.4}$$

Observe that the kth root of any real number is defined as long as $k \neq 0$, as then $1/k = 1/0$, which isn't defined. Therefore, our proposition holds.

2.

$$1/\ln n < \epsilon \tag{1.5}$$
$$1 < \epsilon \ln n \tag{1.6}$$
$$1/\epsilon < \ln n \tag{1.7}$$
$$e^{1/\epsilon} < n \tag{1.8}$$

Therefore, if $N = \lceil e^{1/\epsilon} \rceil$, then this will means for all $n > N$, $|1/\ln n - 0| < \epsilon$, hence our proposition holds.

Proposition 1.1.3

Suppose $a_n \to a$ and $b_n \to b$ as $n \to \infty$ then, as $n \to \infty$:

1. $\lambda a_n \to \lambda a$.
2. $a_n + b_n \to a + b$
3. $a_n - b_n \to a - b$
4. $a_n b_n \to ab$

5. *$a_n/b_n \to a/b$, provided that for all n, $b_n \neq 0$, and $b \neq 0$.*

Proof:

1. To find an N, such that for all $n > N$, $|\lambda a_n - \lambda a| < \epsilon$,for some ϵ, find N such that $|a_n - a| < \epsilon/|\lambda|$, which implies $|\lambda||a_n - a| < \epsilon$, which implies $|\lambda a_n - \lambda a| < \epsilon$.
2. To find an N, such that for all $n > N$, such that $|a_n + b_n - a - b| < \epsilon$, find N_1 such that $|a_n - a| < \epsilon/2$ and N_2 such that $|b_n - b| < \epsilon/2$, and let $N = \max\{N_1, N_2\}$, so then $|a_n - a| + |b_n - b| < \epsilon$, so $|a_n + b_n - a - b| < \epsilon$.
3. Corollary of above.

Proposition 1.1.4 ***(Sandwich Theorem)***

If l_n and u_n are sequences, such that $l_n \to l$ and $u_n \to l$, as $n \to \infty$ and for some $n > N$, $l_n \leq a_n \leq u_n$ for some a_n, then $a_n \to l$, as $n \to \infty$.

Proof: Let N_1 be a positive integer such that for all $n > N$, $|l_n - l| < \epsilon$. And let N_2 be a positive integer such that for all $n > N$, $|u_n - l| < \epsilon$. Let N_3 be a positive integer such that for all $n > N$, $l_n \leq a_n \leq u_n$. Then let $N = \max\{N_1, N_2, N_3\}$. Since $|l_n - l| < \epsilon$, this means $-\epsilon < l_n - l$, i.e. $l - \epsilon < l_n$. Similarly, since $|u_n - l| < \epsilon$, $u_n - l < \epsilon$, so $u_n < l + \epsilon$. This means $l - \epsilon < l_n \leq a_n \leq u_n < l + \epsilon$. This means $l - \epsilon < a_n < l + \epsilon$. $l - \epsilon < a_n$, leads to $-\epsilon < a_n - l$. Similarly, $a_n < l + \epsilon$, leads to $a_n - l < \epsilon$. This means $|a_n - l| < \epsilon$.

For instance, consider $\sin n/n$, this can be bound between $-1/n$ and $1/n$, since the minimum and maximum values of $\sin n$ are -1 and 1 respectively. Both converge to 0, so $\sin n/n$ converges to 0.

Proposition 1.1.5

1. *Ratio Convergence Test: If $\left|\frac{a_{n+1}}{a_n}\right| \leq c < 1$, for some $c \in \mathbb{R}$ and all $n \geq N$ for some positive integer N, then $a_n \to 0$, as $n \to \infty$.*
2. *Ratio Divergence Test: If $\left|\frac{a_{n+1}}{a_n}\right| \geq c > 1$, for some $c \in \mathbb{R}$ and all $n \geq N$ for some positive integer N, then a_n diverges.*
3. *Limit Ratio Test: If $r = \lim_{n\to\infty} |a_{n+1}/a_n|$, then $r < 1$ implies a_n converges, and $r > 1$ implies a_n diverges.*

Proposition 1.1.6

1. *For all c, the sequence $a_n = 1/c^n$ converges to 0, where $|c| > 1$.*
2. *For all c, the sequence $a_n = c^n$, diverges, where $|c| > 1$.*
3. *$1/n!$ converges to 0.*

Proof:

1. The ratio is $1/c$, and if $|c| > 1$, then $1/|c| < 1$ so $|1/c| < 1$, then by the ratio convergence test, the sequence converges to 0.
2. The ratio is c, and if $c > 1$, then by the ratio divergence, test the sequence diverges.
3. The ratio is n, and $n \geq 2$, for all $n > N$, where $N = 1$. So the sequence diverges.

1.2 Series

Definition 1.2.1

Let a_n be a sequence:

1. *Its nth partial sum, S_n is defined as:*

$$S_n = \sum_{n=1}^{n} a_n$$

2. *Its series, S is defined as the sequence of its partial sums, denoted as:*

$$S = \sum_{n=1}^{\infty} a_n$$

Proposition 1.2.1 *For a series to converge, its sequence must converge to 0.*

Definition 1.2.2

- *A geometric series is the series defined over the sequence c^n, for some c.*
- *The harmonic series is the series defined over the sequence $1/n$.*

Proposition 1.2.2 *The geometric series $\sum x^n$ converges to $x/(1-x)$, provided $|x| < 1$.*

Proof: This can be shown by showing that the for nth partial sum:

$$G_n + x^{n+1} = x + xG_n \tag{1.9}$$
$$G_n(1-x) = x - x^{n+1} \tag{1.10}$$
$$G_n = \frac{x - x^{n+1}}{1-x} \tag{1.11}$$

Observe, the limit of the numerator is x assuming $|x| < 1$, and the limit of the denominator is $1 - x$, so the limit of the partial sums, and therefore the series, is $x/(1-x)$.

Proposition 1.2.3 *The harmonic series diverges to infinity.*

Proof: Observe:

$$S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots \quad (1.12)$$

$$= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots \quad (1.13)$$

$$> 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots \quad (1.14)$$

This can be formally, shown as follows:

$$S_{2^n} \geq 1 + \frac{n}{2}$$

This can be shown by induction. Assume true for $n = k$, i.e:

$$S_{2^k} \geq 1 + \frac{k}{2}$$

Then observe, for $2^k \leq n \leq 2^{k+1}$, $1/n > 1/2^{k+1}$, also observe there are 2^k numbers between 2^k and 2^{k+1}, inclusive, so $\sum_{n=2^k}^{2^{k+1}} 1/n > 2^k 1/2^{k+1} = 1/2$. So:

$$S_{2^k+1} > 1 + \frac{k+1}{2}$$

The base case is where $k = 0$. Hence, the harmonic series must diverge, since partial sums can be found that are arbitrarily large.

Proposition 1.2.4 *The series* $\sum 1/n^2$ *converges.*

Proof: To show this, consider:

$$\sum_{n=1}^{\infty} 1/n(n+1)$$

Partial fractions can be used to show that $1/n(n+1) = 1/n - 1/(i+1)$, thus:

$$\sum_{n=1}^{n} \frac{1}{n(n+1)} = 1 - \frac{1}{n+1}$$

Hence, the the series $\sum 1/n(n+1)$ is convergent. It can also be noted that:

$$\sum_{n=2}^{n} \frac{1}{n(n-1)} = \sum_{n=1}^{n-1} \frac{1}{n(n+1)}$$

So this series, must also converge.

Note that, for $n \geq 2$:

$$\frac{1}{n(n-1)} \leq \frac{1}{n} \leq \frac{1}{n(n+1)}$$

So:

$$\sum_{2}^{n} \frac{1}{n(n-1)} \leq \sum_{2}^{n} \frac{1}{n} \leq \sum_{2}^{n} \frac{1}{n(n+1)}$$

Note, that the sequence of partial sums $\sum_2^n 1/n^2$ are bound between two convergent sequences of partial sums, so $\sum_2^\infty 1/n^2$ must converge by the sandwich theorem, therefore $\sum_1^\infty 1/n^2$ also converges.

Proposition 1.2.5 *The series below, where P is the set of prime numbers is divergent:*

$$\sum_{p \in P} \frac{1}{p}$$

Proposition 1.2.6

Assume $\sum c_i$ is a convergent series and $\sum d_i$ is a divergent series.

1. *Comparison Test: If $a_n \leq \lambda c_n$, for all $n > N$, where N is a positive integer, then $\sum a_i$ converges, if a_n is non-negative.*
2. *Comparison Test: If $a_n \geq \lambda d_n$, for all $n > N$, where N is a positive integer, then $\sum a_i$ diverges, if a_n is non-negative.*
3. *Limit Comparison Test: If $\lim_{n\to\infty} a_n/c_n$ exists then $\sum a_n$ converges, if a_n is non-negative.*
4. *Limit Comparison Test: If $\lim_{n\to\infty} d_n/a_n$ exists then $\sum a_n$ converges, if a_n is non-negative.*
5. *If $|a_{n+1}/a_n| \geq 1$ for all n, where $n > N$, then $\sum a_n$ diverges, if a_n is non-negative.*
6. *If $|a_{n+1}/a_n| \leq k < 1$ for all n, where $n > N$, then $\sum a_n$ converges, if a_n is non-negative.*
7. *If $r = \lim_{n\to\infty} |a_{n+1}/a_n|$, $r < 1$, implies $\sum a_n$ converges, and $r > 1$, implies $\sum a_n$ diverges.*

Proposition 1.2.7 *If $\sum |a_n|$ converges then so does $\sum a_n$.*

Suppose $a_n = f(n)$, where $f(x)$ is a strictly decreasing function. Then:

$$\sum_{n=1}^{\infty} a_{n+1} < \int_1^{\infty} f(x)\ dx < \sum_{n=1}^{\infty} a_n$$

Therefore, if the integral diverges, by the right-hand inequality your series diverges. If the integral converges, by the left-hand inequality your series converges. This is called the integral test.

1.3 Power Series

Proposition 1.3.1

1. *The derivative of* x_n *with respect to* x *is* nx^{n-1}*.*
2. *The derivative of* $\ln x$ *with respect to* x *is* $1/x$*.*
3. *The derivative of* $\sin x$*,* $\cos x$ *and* $\tan x$ *is* $\cos x$*,* $-\sin x$ *and* $\sec^2 x$ *respectively.*
4. *The derivative of* $f(g(x))$*, with respect to* x*, is* $f'(g(x))g'(x)$*, where* f' *and* g' *are the derivatives with respect to* x *of* f *and* g *respectively.*
5. *The derivative of* e^x *with respect to* x *is* e^x*, and the derivative of* a^x *is* $a^x \ln a$*.*

Proposition 1.3.2

Let u *and* v *be functions of* x*, and* u' *and* v' *be their respective derivatives, with respect to* x*.*

1. *The derivative of* λu*, where* λ *is a constant is* $\lambda u'$*.*
2. *The derivative of* $u + v$ *with respect to* x *is* $u' + v'$*.*
3. *The derivative of* uv *with respect to* x *is* $uv' + u'v$*.*
4. *The derivative of* u/v *with respect t* x *is* $(vu' - uv')/v^2$*.*
5. $\int uv'\ dx = uv - \int u'v\ dx$*.*

Definition 1.3.1

1. *A power series is a way of writing a function,* $f(x)$ *in the form:*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

2. *The radius of convergence of a power series is the largest* $r \in \mathbb{R}$*, such that* $|x| < r$ *defines a set of* x *for which a power series converges. The radius of convergence can also be defined around a point, such that* $|x - a| < r$ *defines a set of* x *around* a *for which the power series converges.*

Definition 1.3.2

Suppose $f(x)$ is a function, that is differentiable infinitely many times.

1. *Its Maclaurin Series expansion is a power series defined as:*

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!}$$

   Where $f^{(n)}(x)$ is the nth derivative of $f(x)$.

2. *Its Taylor Series expansion about some $a \in \mathbb{R}$ is defined as:*

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(a)\frac{(x-a)^n}{n!}$$

In order, to compute a Taylor series, it is usually truncated at a certain point. It should be noted that:

$$f(x) = \sum_{n=0}^{k} f^{(n)}(a)\frac{(x-a)^n}{n!} + f^{(k+1)}(c)\frac{(x-a)^{k+1}}{(k+1)!}$$

The final term is called the Lagrange error term, and it is the case that $x < c < a$, or $a < c < x$. However, since it is difficult to find c, the bounds can be used to generate a worst-case error term. A different form of the error term is the Cauchy error term, which is given as:

$$f(x) - \sum_{n=0}^{k} f^{(n)}(a)\frac{(x-a)^n}{n!}$$

Power series can be used to solve differential equations. This can be done by replacing $f(x)$, $f'(x)$, etc. in terms of a power series, and then equating coefficients.

1.4 Abstract Algebra

Definition 1.4.1

A group, $(G, \star)$ consists of a set G, and a binary operation $*$, where:

- *Closure:* $\forall a, b \in G \;\; a \star b \in G$
- *Associativity:* $\forall a, b, c \in G \;\; (a \star b) \star c = a \star (b \star c)$
- *Identity:* $\exists e \in G \;\; \forall a \in G \;\; e \star a = e = a \star e$

- *Inverses:* $\forall a \in G \;\; \exists \tilde{a} \in G \;\; a \star \tilde{a} = e = \tilde{a} \star a$

Proposition 1.4.1

1. *The set G is non-empty,* $G \neq \varnothing$.
2. *Left cancellation property: If* $a * b = a * c$*, then* $b = c$.
3. *Right cancellation property: If* $b * a = c * a$*, then* $b = a$.
4. *There is one identity element i.e.* $\forall e_1, e_2 \in G$*, where* $\forall a \;\; e_1 \star a = a$ *and* $\forall a \;\; e_2 \star a = a$*, then* $e_1 = e_2$.
5. *An element has one inverse.*
6. *Two distinct elements do not share an inverse.*

Proof:

1. This is as $e \in G$.
2. This is shown by left multiplying by $\tilde{a}$.
3. This is shown by right multiplying by $\tilde{a}$.
4. This follows from the previous propositions.
5. This follows from the previous propositions.
6. This follows from the previous propositions.

Proposition 1.4.2 *The inverse of* $(x \star y)$ *is* $\tilde{y} \star \tilde{x}$.

Proof: $(\tilde{y} \star \tilde{x}) \star (x \star y) = \tilde{y} \star y = e$

Definition 1.4.2 *An abelian group* $(G, \star)$*, is a group where* $a \star b = b \star a$.

Examples of groups include the integers under addition i.e. $(\mathbb{Z}, +)$. The same applies for $(\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$. Also observe that the following are groups: $(\mathbb{Q} \setminus \{0\}, \times)$, $(\mathbb{R} \setminus \{0\}, \times)$.

1.5 Vector Space

Definition 1.5.1 *A real-valued vector space is a set* V*, with two operations:*

$$+ : V \times V \to V \text{ (Vector addition)}$$

$$* : \mathbb{R} \times V \to V \text{ (Scalar multiplication)}$$

Such that:

- *$(V, +)$ forms an abelian group.*
- $(ab)\mathbf{c} = a(b\mathbf{c})$
- *Scalar multiplication has an identity element.*
- *Distributivity of scalar multiplication over vector addition i.e. : $a(\mathbf{b} + \mathbf{c}) = a\mathbf{b} + a\mathbf{c}$.*
- *Distributivity of scalar multiplication over field addition i.e. : $(a + b)\mathbf{c} + a\mathbf{c} + b\mathbf{c}$.*

Definition 1.5.2

- *A linear combination is a finite set overs vectors, $v_1, \ldots v_n \in V$ with corresponding constants: $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ is:*

$$\sum_{i=1}^{n} \lambda_i v_i$$

- *A set of vectors $v_1, \ldots, v_n$ is linearly independent if there exists a linear combination, where it is not the case that $\lambda_1 = \ldots = \lambda_n = 0$, such that:*

$$\sum_{i=1}^{n} \lambda_i v_i = 0$$

- *The span of a set of vectors is the set of all vectors that can be expressed as linear combination of these vectors.*
- *A generating set of a vector space is a set of vectors the span a vector space.*
- *A basis of a vector space is a generating set that is minimal.*
- *The dimension of a vector space is the cardinality of a basis of a vector space.*
- *A subspace of a vector space V is a set of vectors V' such that V' is non-empty and V' is closed under vector addition and scalar multiplication.*
- *A mapping ϕ between two vector spaces, V and W is linear if:*

$$\forall v, w \in V,\ \lambda \in F\ \phi(v + \lambda w) = \phi(v) + \lambda\phi(w)$$

- *A homomorphism is a linear mapping.*
- *An endomorphism is a homomorphism between the same vector space.*
- *An automorphism is a bijective endomorphism.*
- *The identity mapping of a vector space V, id_V is defined as $id_V : x \mapsto x$.*

Proposition 1.5.1

1. *If there are two different linear combinations over the same vectors that equal the same value, then the linear combination is not linearly independent.*

2. *If a linear combination over $v_1, \ldots, v_n$ equals some vector $v*$, then the set $v_1, \ldots, v_n, v*$ is linearly dependent.*
3. *If the v' is in the span of the generating set of $v_1, \ldots, v_n$, then the span of the generating set $v_1, \ldots, v_n, v'$ is equal to the original span.*
4. *A basis is linearly independent.*
5. *The dimension of a vector space is unique.*

Note that an empty generating set spans the 0 vector, since a linear combination of no vectors equals the 0 vector.

1.6 Matrices

Definition 1.6.1 *If $X \in R^{n,m}$, where $m, n \in \mathbb{N}_0$, where $m \times n$ is the dimension of a real-valued matrix, where m is the number of rows and n is the number of columns.*

This can be denoted as below:

$$X = [x_{ij}] = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \ldots & x_{1m} \\ x_{21} & x_{22} & x_{23} & \ldots & x_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \ldots & x_{nm} \end{bmatrix}$$

The shorthand $[x_{ij}]$ is used as a short-hand to refer generally to elements within matrices.

A matrix in the form $R^{n,1}$ is a row vector and a matrix in the form $R^{1,n}$ is a column vector.

Let $A \in R^{n,m}$, then let $\text{row}(A, k)$, where $1 \leq k \leq n$, be a row vector from the kth row of A. Similarly let $\text{col}(A, k)$, where $1 \leq k \leq m$, be a row vector from the kth row of A.

Addition over two matrices $A, B \in R^{n,m}$ can be defined as follows, where $A + B = C \in R^{n,m}$:

$$[c_{ij}] = [a_{ij} + b_{ij}]$$

Scalar multiplication of a matrix $A \in R^{n,m}$ with scalar $\lambda \in R$ can be defined as follows, where $C = \lambda A$:

$$[c_{ij}] = [\lambda a_{ij}]$$

Definition 1.6.2 *A square matrix is a matrix of dimensions $n \times n$.*

Definition 1.6.3 *Matrix multiplication is defined as* $* : R^{n,m} \times R^{m,l} \to R^{n,l}$*. The result of the multiplication of* $A \in R^{n,m}$ *and* $B \in R^{m,l}$ *is* $C \in R^{n,l}$*, where:*

$$c_{ij} = \sum_{k=1}^{m} a_{ik} * b_{kj}$$

This means:

$$\operatorname{col}(C, k) = \sum_{i=1}^{m} b_{ik} \operatorname{col}(A, i)$$

In other words, each column vector in C is a linear combination of the column vectors in A defined by the corresponding column in B. Similarly:

$$\operatorname{row}(C, k) = \sum_{i=1}^{m} a_{ki} \operatorname{row}(B, i)$$

I.e. each row in C is a linear combination of the rows in B defined by the corresponding row in A.

Proposition 1.6.1

1. *Matrices in* $R^{n,m}$ *form a vector space.*
2. *Matrix multiplication is associative.*

Definition 1.6.4

- *The Kronecker delta function is defined as follows:*

$$\delta_{ij} = \begin{cases} 0, & \textit{if } i \neq j \\ 1, & \textit{if } i = j \end{cases}$$

- *The identity matrix* $I_n = [\delta_{ij}]$ *is an* $R^{n,n}$ *matrix.*
- *The identity matrix* I_n *is the identity for matrix multiplication with matrices in* $R^{n,n}$*.*
- *The rank of a matrix* A*, denoted as* $\operatorname{rk}(a)$ *is the number of linearly independent rows it has, which is equal to the number of linearly independent columns its has.*
- *A matrix* $A^{n,n}$ *is invertible if there exists a matrix* $A^{-1} \in R^{n,n}$*, such that* $AA^{-1} = I_n = AA^{-1}$*.*
- *The set of invertible matrices in* $R^{n,n}$ *forms a group,* $GL_2(\mathbb{R})$*.*
- *If a* $R^{n,m}$ *matrix has rank* $\min\{n, m\}$*, it is considered full-rank.*
- *A matrix which isn't full rank is rank deficient.u*

- *A matrix A is upper triangular if $a_{ij} = 0$ when $i > j$.*
- *A matrix A is lower triangular if $a_{ij} = 0$ when $i < j$.*
- *A matrix A is diagonal if it is lower triangular and upper triangular. This means $a_{ij} = 0$, when $i \neq j$.*
- *A matrix $A \in R^{n,n}$ is similar to a matrix $\tilde{A} \in R^{n,n}$ if there exists an invertible matrix $S \in R^{n,n}$ such that $\tilde{A} = S^{-1}AS$.*
- *A matrix is diagonalisable if it is similar to a diagonal matrix.*
- *The row space of a matrix is the span of its row vectors.*
- *The column space of a matrix is the span of its column vectors.*
- *For any matrix A, its kernel $\ker(A)$, is the set: $\{x \mid Ax = 0\}$.*
- *For any matrix $A \in R^{n,m}$, its image, $\mathrm{im}(A)$, is the set: $\{Ax \mid x \in R^{m,n}\}$.*
- *A matrix that is not invertible is singular or degenerate.*
- *A matrix $A \in \mathbb{R}^{n,m}$ is equivalent to $\tilde{A}\mathbb{R}^{n,n}$, if there exist two invertible matrices, $S \in \mathbb{R}^{m,m}$ and $T \in \mathbb{R}^{n,n}$ such that $\tilde{A} = T^{-1}AS$.*

Proposition 1.6.2 *The rank-nullity theorem. For any matrix $A \in R^{n,m}$*

$$\dim(\mathrm{im}(A)) + \dim(\ker(A)) = m$$

Proposition 1.6.3

The following are equivalent for a matrix $A \in R^{n,n}$:

1. *A matrix is has full-rank.*
2. *A matrix is invertible.*
3. *The kernel has dimension 0.*
4. *The kernel is equal to $\{0\}$.*
5. *The image equals $R^{n,n}$.*
6. *The dimension of the image is n.*

1.7 Transposes

Definition 1.7.1

- *The transpose of a matrix A, where $A \in R^{n,m}$ denoted as A^T, where $A^T \in R^{m,n}$ is defined:*

$$[a_{ij}]^T = [a_{ji}]$$

- *A matrix $A \in R^{n,n}$ is symmetric if $A = A^T$.*

Proposition 1.7.1

1. $A^{T^T} = A$
2. $(A + B)^T = A^T + B^T$.
3. $(AB)^T = B^T A^T$
4. *If A is invertible, $(A^T)^{-1} = (A^{-1})^T$, where $A \in R^{n,n}$*
5. $\mathrm{rk}(A) = \mathrm{rk}(A^T)$

1.8 Gaussian Elimination

Definition 1.8.1 *The leading entry of a row if the first non-zero entry in a row.*

Definition 1.8.2 *A matrix is in row echelon form if:*

- *The leading entry, if exists, of every row is 1.*
- *The entries below a leading entry are 0.*
- *Rows that do not have a leading entry, i.e. entirely consist of 0s, are below rows with a leading entry.*

Proposition 1.8.1 *A matrix $A \in R^{n,m}$ in row echelon form is upper triangular.*

Definition 1.8.3 *A matrix is in reduced row echelon form if it is in row echelon form, and every entry above a leading entry is 0.*

Proposition 1.8.2 *Every matrix $A \in K^{n,m}$ can be converted to row echelon form.*

Algorithm:

1. Locate the first column i containing a non-zero entry. If none can be found, then the matrix is already in row echelon form.
2. Swap the row containing this entry with the first row.
3. Suppose the value of this entry is a. Divide all items in the first row by a^{-1}. This means this entry is now 1.
4. Now, for every non-zero entry in column i, in position (k, j), subtract a_{1j} times row 1 from row k. This means that all values below a_{1j} will be 0.
5. The matrix is now in row echelon form in column i and row 1. If column $i = 1$, then the matrix looks like this:

$$A = \left[\begin{array}{c|c} 1 & \star \\ \hline 0 & A' \end{array}\right]$$

Where A' is the block matrix that is not in row echelon form. If there are zero columns, before the first column, the matrix also looks like this, except with additional columns full of zero to the left.

6. Now recursively apply the Gaussian Elimination algorithm to A'. This will preserve the row echelon form of row 1 and column 1. This is as multiplying any rows below row 1 by λ will not change values in column 1, as they are all 0. Similarly subtracting row c from d, where $c, d > 1$ won't change the values in column 1, as the $a_{c1} = 0$.

Proposition 1.8.3 *Every matrix can be converted to reduced row echelon form.*

Algorithm:

- For each non-zero value a in row j above a leading entry in row i, subtract a times row i from row j. This won't affect any leading entries as $j < i$, and all values to the left of the leading entry are 0.

Problem 1 *Solve a inhomogeneous equation system $Ax = 0$.*

Algorithm:

To solve an equation system $Ax = 0$, perform Gaussian elimination. Then find the basic variables in terms of the free variables, and that vector is your solution. If there are no free variables, then there is 1 solution. If there are free variables, there are infinitely many solutions.

Problem 2 *Solve an inhomogeneous linear equation system $Ax = b$*

Algorithm:

To solve an equation system $Ax = b$, perform Gaussian elimination on the augmented matrix $[A|b]$. If there are zero rows, and the corresponding values in b are not 0, then there are no solutions. Then substitute arbitrary values for the free variables, and then get values for the basic variables, and you have obtained the particular solution. Add this to the solution for $Ax = 0$, and then you have a general solution.

Problem 3 *Find the inverse of $A \in R^{n,n}$.*

Algorithm:

Solve the inhomogeneous linear equation system $Ax = I_n$.

Problem 4 *Find basis for the span of a set of vectors.*

Algorithm: Perform Gaussian elimination on a matrix whose column vectors are these vectors, and take the column vectors corresponding to the pivot columns as the basis vectors.

Problem 5 *Find a simple basis for a set spanned by some vectors*

Algorithm: Perform Gaussian elimination on a matrix whose row vectors are these vectors, and take the non-zero rows as the basis vectors.

1.9 Determinants

Proposition 1.9.1

1. *The determinant of a matrix $A \in R^{2,2}$ is $ad - bc$, where:*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

2. *If a matrix $A \in R^{n,n}$ is upper or lower triangular then* $\det(A) = \prod_{i=1}^{n} a_{ii}$.
3. *For a matrix $A \in R^{n,n}$, where, $n \geq 1$ it is the case that for any column $1 \leq k \leq n$:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} \det\left(\tilde{A}_{ik}\right)$$

   Or row $1 \leq k \leq n$:

$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} \det\left(\tilde{A}_{ki}\right)$$

   Where $\tilde{A}_i k$ is the matrix A with the column k and row i removed.
4. $\det(AB) = \det(A)\det(B)$.
5. *If A is an invertible matrix, then* $\det(A^{-1}) = 1/\det(A)$.
6. *If A is an invertible matrix, then* $\det\left(A^T\right) = \det(A)$.
7. $\det(A) = 0$ *if and only if A is singular.*
8. *Similar matrices have the same determinant.*

1.10 Transformations

Definition 1.10.1 *An ordered basis is a basis, where each vector is given a number, such that vectors in the vector space represented by the basis can be written as co-ordinates,* $(\lambda_1, \ldots, \lambda_n)$*, where the value of the expression written as co-ordinates is the value of the linear combination* $\lambda_1, \ldots, \lambda_n$*over the basis vectors.*

Problem 6 *Represent a homomorphism between two vector spaces* V *with basis* $\tilde{V}$ *and* W *and basis* $\tilde{W}$*,* ϕ *as a transformation matrix.*

Algorithm: Take the ordered basis vectors of V in terms of $\tilde{V}$ and find their after being transformed by ϕ in terms of the ordered basis of W, $\tilde{W}$. Then putting them as the column vectors in order in a matrix, will give you the transformation matrix.

Problem 7 *Change basis from basis* B *to basis* B'*.*

Algorithm: Follow the same algorithm as a homomorphism between two vector spaces, except it is between the same vector space, just with different bases.

Problem 8 *Suppose you have an endomorphism* ϕ *that operates on basis* B*, and you want it to operate on basis* C*.*

Algorithm: Let T be the transformation matrix that goes from B to C, then $\phi = T\phi' T^{-1}$, so $\phi' = T^{-1}\phi T$.

Definition 1.10.2

An affine transformation is a mapping in the form, where f' *is a homomorphism:*

$$f(x) = f'(x) + y$$

An affine subspace is a subspace made from the span of vectors $v_1, \ldots, v_n$ *and a translation vector* v'*:*

$$v' + [v_1, \ldots, v_n]$$

1.11 Intersections

Problem 9 *Find the intersection of two subspaces* B *and* C*, with bases* $b_1, \ldots, b_n$*, and* $c_1, \ldots c_m$*.*

Algorithm: You're trying to find the vectors that satisfy:

$$\sum_{i=1}^{n} \lambda_i b_i - \sum_{i=1}^{m} \psi_i c_i = 0$$

This means, that it should satisfy a homogenous linear equation system $Ax = 0$, where A is made from the basis vectors of B and the negated basis vectors of C. This can be done by performing Gaussian elimination. Note, you only need to find either $\lambda_1, \ldots, \lambda_n$ or $\psi_1, \ldots, \psi_n$.

Problem 10 *Find the intersection of two affine subspaces B and C, with bases $b_1, \ldots, b_n$, and $c_1, \ldots c_m$, and translation vectors b' and c'. Note, you only need to find either $\lambda_1, \ldots, \lambda_n$ or $\psi_1, \ldots, \psi_n$.*

Algorithm: You're trying to find the vectors that satisfy:

$$b' + \sum_{i=1}^{n} \lambda_i b_i - c' - \sum_{i=1}^{m} \psi_i c_i = 0$$

$$\sum_{i=1}^{n} \lambda_i b_i - \sum_{i=1}^{m} \psi_i c_i = c' - b'$$

This means, that it should satisfy a homogenous linear equation system $Ax = c' - b'$, where A is made from the basis vectors of B and the negated basis vectors of C. This can be done by performing Gaussian elimination.

1.12 Eigenvalues

Definition 1.12.1

- *For an endomorphism $\phi : V \to V$ and a scalar $\lambda \in \mathbb{R}$, λ is an eigenvalue of ϕ if here exists $x \in V \setminus \{0\}$ such that:*

  $$\phi(x) = \lambda x$$

  x is the corresponding eigenvector of the eigenvalue of λ.

- *The set of all eigenvectors of ϕ with respect to λ is the eigenspace of ϕ with respect to λ.*
- *The spectrum of ϕ is the set of all eigenvalues.*
- *The characteristic polynomial is of matrix $A \in R^{n,n}$, is $\det(A - \lambda I_n)$, and the roots of $\det(A - \lambda I_n) = 0$ are the eigenvalues of A.*
- *The coefficient for the lowest power of the characteristic polynomial is equal to $\det(A)$ and the coefficient for the second highest power is equal to $(-1)^{n-1}\operatorname{tr}(A)$, where $\operatorname{tr}(A)$ is the sum of the diagonal elements. The coefficient for the highest power is $(-1)^n$.*
- *The algebraic multiplicity of an eigenvalue is the number of times it is repeated as a root in the characteristic polynomial.*
- *The geometric multiplicity of an eigenvalue λ is the dimension of the eigenspace λ.*
- *The geometric multiplicity cannot be greater than the algebraic multiplicity.*

- *Geometric multiplicity is be at least 1.*
- *If the spectrum is non-empty, that means* $(A - \lambda I_n)x = 0$ *can be solved where* $x \neq 0$, *therefore,* $(A - \lambda I_n)$ *is singular, and is rank deficient.*
- *The eigenbasis is a basis of* V *formed from the eigenvectors of* ϕ.

Proposition 1.12.1

- *Similar matrices have the same eigenvalues.*
- *A matrix* $A \in \mathbb{R}^{n,n}$ *is only diagonalisable if for an eigenvalue* λ *which has algebraic multiplicity* r *it is the case that* $\dim(\mathrm{im}(\phi - \lambda\,\mathrm{id}_V)) = n - r$.
- *This means that* $\dim(\ker(\phi - \lambda\,\mathrm{id}_V)) = r$.
- *This means that algebraic multiplicity equals geometric multiplicity.*
- *It also means that the sum of the geometric multiplicities equals* n.
- *And there exists a basis of in* V *consisting of the eigenvectors of the endomorphism.*
- *Suppose for a matrix* $A \in R^{n,n}$ *with* n *eigenvalues, then eigenvalue* λ_i *has algebraic multiplicity* r_i *and is diagonalisable, then the diagonal matrix such that* $A = SDS^{-1}$, *is:*

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

  Note that in the above matrix λ_i *is repeated* r_i *times. Also note that it follows* $D = S^{-1}AS$.

- *S is a change of basis matrix converting a vector in the standard basis to the eigenbasis, and, therefore,* S^{-1} *is a change of basis matrix converting a vector in the eigenbasis to the standard basis. S is formed by combining an eigenbasis of* A, *such that the length of each vector is 1.*

1.13 Rotations

A rotation counter-clockwise in $\mathbb{R}^2$ is represented by the matrix:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

In $\mathbb{R}^3$, a counter-clockwise rotation is a defined around an axis, which means it is carried as if facing the other two axes from the end of the axis it is around.

$$[e_1|e_2|e_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Counter-clockwise rotation of θ around e_1 is defined as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

Counter-clockwise rotation of θ around e_2 is defined as:

$$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

Counter-clockwise rotation of θ around e_3 is defined as:

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

To generalise rotations to higher dimensions, we think of it as keeping $n-2$ dimensions fixed. The dimensions that aren't fixed are $1 \leq i < j \leq n$. The rotation matrix is given be:

$$\begin{bmatrix} I_{i-1} & 0 & \dots & \dots & 0 \\ 0 & \cos\theta & 0 & -\sin\theta & 0 \\ 0 & 0 & I_{j-i} & 0 & 0 \\ 0 & \sin\theta & 0 & \cos\theta & 0 \\ 0 & \dots & \dots & 0 & I_{n-j} \end{bmatrix}$$

Proposition 1.13.1

Suppose $R(\theta)$ defines a rotation matrix:

- $R(\theta)R(\phi) = R(\theta + \phi)$
- *Rotation matrices have no real eigenvalues unless $\theta = n\pi$, where $n \in \mathbb{Z}$.*
- *Rotations preserve lengths and distances i.e. $|R(\theta)x| = |x|$ and $|R(\theta)x - R(\theta)y| = |x - y|$.*
- *Rotations in 2 dimensions are commutative but that is not generally the case for 3 or more dimensions.*