Chapter 1

Linear Algebra

Definition 1.1

A vector space (or linear space) V over a field $\mathbb{F}$ consists of a set on which two operations (called addition and multiplication respectively here) are defined so that;

(A) (V is Closed Under Addition) For all $\mathbf{x}, \mathbf{y} \in V$, there exists a unique element $\mathbf{x} + \mathbf{y} \in V$.

(M) (V is Closed Under Scalar Multiplication) For all elements $a \in \mathbb{F}$ and elements $\mathbf{x} \in V$, there exists a unique element $a\mathbf{x} \in V$.

Such that the following properties hold:

(VS 1) (Commutativity of Addition) For all $\mathbf{x}, \mathbf{y} \in V$, we have $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

(VS 2) (Associativity of Addition) For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, we have $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.

(VS 3) (Existance of The Zero/Null Vector) There exists an element in V denoted by $\mathbf{0}$, such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in V$.

(VS 4) (Existance of Additive Inverses) For all elements $\mathbf{x} \in V$, there exists an element $\mathbf{y} \in V$ such that $\mathbf{x} + \mathbf{y} = \mathbf{0}$.

(VS 5) (Multiplicative Identity) For all elements $x \in V$, we have $1\mathbf{x} = \mathbf{x}$, where 1 denotes the multiplicative identity in $\mathbb{F}$.

(VS 6) (Compatibility of Scalar Multiplication with Field Multiplication) For all elements $a, b \in \mathbb{F}$ and elements $\mathbf{x} \in V$, we have $(ab)\mathbf{x} = a(b\mathbf{x})$.

(VS 7) (Distributivity of Scalar Multiplication over Vector Addition) For all elements $a \in \mathbb{F}$ and elements $\mathbf{x}, \mathbf{y} \in V$, we have $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$.

(VS 8) (Distributivity of Scalar Multiplication over Field Addition) For all elements $a, b \in \mathbb{F}$, and elements $\mathbf{x} \in V$, we have $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$.

Theorem 1.2

Let V be a vector space and W a subset of V. Then W is a subspace of V iff the following 3 conditions hold for the operations defined in V.

(a) $\mathbf{0} \in W$

(b) $\mathbf{x} + \mathbf{y} \in W$ whenever $\mathbf{x} \in W$ and $\mathbf{y} \in W$.

(c) $c\mathbf{x} \in W$ whenever $c \in \mathbb{F}$ and $\mathbf{x} \in W$.

Definition 1.3

A subset S of a vector space V *generates* (or *spans*) V iff $\text{span}(S) = V$. In this case, we also say that the vectors of S generate (or span) V.

Definition 1.4

Let V be a vector space and S a nonempty subset of V. A vector $v \in V$ is called a *linear combination* of vectors of S iff there exists a finite number of vectors $u_1, u_2, \ldots, u_n$ in S and scalars $a_1, a_2, \ldots, a_n$ in $\mathbb{F}$ such that

$$v = \sum_{i=1}^{n} a_i u_i.$$

In this case we also say that v is a linear combination of $u_1, u_2, \ldots, u_n$ and call $a_1, a_2, \ldots, a_n$ the *coefficients* of the linear combination

Definition 1.5

A set subset S of a vector space V is called *linearly dependent* iff there exists a finite number of distinct vectors $u_1, u_2, \ldots, u_n$in S and scalars $a_1, a_2, \ldots, a_n$ not all zero, such that

$$a_1u_1 + a_2u_2 + a_nu_n = \mathbf{0}.$$

Definition 1.6

A *basis* β for a vector space V is a linearly independent subset of V that generates V. If β is a basis for V, we also say that the vectors of β form a basis for V.

Theorem 1.7: The Rank-Nullity Theorem.

For any vector spaces V and W, and a linear operator $T : V \to W$, it holds that

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

General Information

- Let $\mathbf{A}$ be an $m \times n$ matrix, and $\mathbf{a}_j$ its jth column. For any $\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^\top$,

$$\mathbf{Ax} = \sum_{j=1}^{n} x_j \mathbf{a}_j.$$

- Let $\mathbf{A}$ and $\mathbf{B}$ be matrices having n rows. For any matrix $\mathbf{M}$ with n columns, we have

$$(\mathbf{A} \mid \mathbf{B}) = (\mathbf{MA} \mid \mathbf{MB}).$$

Definition 1.8

A system $\mathbf{Ax} = \mathbf{b}$ is *homogeneous* iff $\mathbf{b} = 0$; otherwise it is *nonhomogeneous*.

Theorem 1.9

For any matrix, its row space, column space, and rank are identical.

Theorem 1.10

A system $\mathbf{Ax} = \mathbf{b}$ of m linear equations in n unknowns has a solution space of dimension $n - \text{rank}(A)$.

Definition 1.11

A system $\mathbf{Ax} = \mathbf{b}$ of linear equations is *consistent* iff its solution set is nonempty; otherwise it is *inconsistent*.

Theorem 1.12: The Rouché-Capelli Theorem.

A system $\mathbf{Ax} = \mathbf{b}$ is consistent iff $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}|\mathbf{b})$.

Definition 1.13

A matrix is said to be in *reduced row echelon form* iff

- Any row containing a nonzero entry precedes any row in which all the entries are zero (if any).
- The first nonzero entry in each row is the only nonzero entry in its column.
- The first nonzero entry in each row is 1 and it occurs in a column to the right of the first nonzero entry in the preceding row.

- Gaussian elimination.
  - In the forward pass, the augmented matrix is transformed into an upper triangular matrix in which the first nonzero entry of each row is 1 and it occurs in a column to the right of the first nonzero entry of each preceding row.
  - In the backward pass, the upper triangular matrix is transformed into reduced row echelon form by making the first nonzero entry of each row the only nonzero entry of its column.
- Gaussian elimination always reduecs a matrix to its rref form.
- Let $\mathbf{A}$ be an invertible $n \times n$ matrix. Then, for some elementary row matrices $\mathbf{E}_1$ to $\mathbf{E}_p$,
$$\mathbf{E}_p\mathbf{E}_{p-1}\dots\mathbf{E}_1(\mathbf{A}\,|\,\mathbf{I}_n) = \mathbf{A}^{-1}(\mathbf{A}\,|\,\mathbf{I}_n) = (\mathbf{I}_n\,|\,\mathbf{A}^{-1}).$$
In other words, we can perform Gaussian elimination, so that $(\mathbf{A}\,|\,\mathbf{I}_n) \to (\mathbf{I}_n\,|\,\mathbf{A}^{-1})$.
- Let $\mathbf{A} := \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{pmatrix}$ be $m \times n$ matrix, and $\mathbf{A}' := \begin{pmatrix}\mathbf{a'}_1 & \mathbf{a'}_2 & \cdots & \mathbf{a'}_n\end{pmatrix}$ its rref. Then, $\{\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \dots, \mathbf{a}_{k_m}\}$ is linearly independent iff $\{\mathbf{a'}_{k_1}, \mathbf{a'}_{k_2}, \dots, \mathbf{a'}_{k_m}\}$ is. Moreover, the row space of $\mathbf{A}$ and $\mathbf{A}'$ are clearly identical.
- Finding a basis for an intersection of subspaces. Let V and W be subspaces of $\mathbb{F}^n$ generated by the columns of the $n \times m$ matrix $\mathbf{A}$ and $n \times k$ matrix $\mathbf{B}$, respectively. Find a basis for the subspace $V \cap W$.
  1. First notice that $\mathbf{v} \in V \cap W$ iff
  $$\mathbf{v} = \mathbf{Ax}_1 = \mathbf{Bx}_2$$
  for some $\mathbf{x}_2 \in \mathbb{F}^m$ and $\mathbf{x}_2 \in \mathbb{F}^k$. That is,
  $$\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}\begin{pmatrix}\mathbf{x_1} \\ -\mathbf{x_2}\end{pmatrix} = \mathbf{0}.$$
  So, equivalently, we write
  $$\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}\mathbf{y} = \mathbf{0}.$$
  for some $\mathbf{y} \in \mathbb{F}^{m+k}$. As such, by row reducing $\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}$, we find a basis
  $$\beta := \left\{\begin{pmatrix}\mathbf{u_1} \\ \mathbf{u'_1}\end{pmatrix}, \begin{pmatrix}\mathbf{u_2} \\ \mathbf{u'_2}\end{pmatrix}, \dots, \begin{pmatrix}\mathbf{u_r} \\ \mathbf{u'_r}\end{pmatrix}\right\},$$
  where $\mathbf{u}_i \in \mathbb{F}^m$ and $\mathbf{u}_i \in \mathbb{F}^k$. Now, a generating set for $V \cap W$ is
  $$\Gamma := \{\mathbf{Au_1}, \mathbf{Au_2}, \dots, \mathbf{Au_r}\}.$$
  Alternatively, another generating set for $V \cap W$ is
  $$\Delta := \{\mathbf{Bu'_1}, \mathbf{Bu'_2}, \dots, \mathbf{Bu'_r}\}.$$
  From here, it is simple to choose bases $\gamma \subseteq \Gamma$ and $\delta \subseteq \Delta$ for $V \cap W$.
  (Naturally, it holds that $\mathbf{Au_i} + \mathbf{Bu'_i} = 0$.)

2. An alternative method. By row reduction, we can calculate

$$\begin{aligned} r := \dim(V \cap W) &= \dim(U) + \dim(V) - \dim(U+V), \\ &= \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B}) - \operatorname{rank}\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}, \\ &= \operatorname{rank}\left(\mathbf{A}^\top\right) + \operatorname{rank}\left(\mathbf{B}^\top\right) - \operatorname{rank}\begin{pmatrix}\mathbf{A}^\top \\ \mathbf{B}^\top\end{pmatrix}. \end{aligned}$$

   Then, a basis for $V \cap W$ can be formed by choosing r linearly independent columns of $\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}$, or rows of $\begin{pmatrix}\mathbf{A}^\top \\ \mathbf{B}^\top\end{pmatrix}$.

3. Another alternative, probably the best option! Skip the row reduction of **A** and **B** in the above method. We just reduce

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix} \to \begin{pmatrix}\mathbf{A}' & \mathbf{B}'\end{pmatrix}.$$

   Let $\mathbf{c_i}$ and $\mathbf{c_i'}$ be the ith column of $\begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}$ and $\begin{pmatrix}\mathbf{A}' & \mathbf{B}'\end{pmatrix}$, respectively. We compare the columns of A' and B' to find (with relative ease) a basis $\beta' := \left\{\mathbf{c'_{i_1}}, \mathbf{c'_{i_2}}, \ldots, \mathbf{c'_{i_r}}\right\}$ for the intersection of the column spaces of A' and B'. Then, $\beta := \left\{\mathbf{c_{i_1}}, \mathbf{c_{i_2}}, \ldots, \mathbf{c_{i_r}}\right\}$ is a basis for $V \cap W$ (the intersection of the column spaces of A and B).

4. A fourth method for when I learn about orthogonal complements.

Definition 1.14

Let $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$. If $n = 1$, so that $A = (a_1 1)$, we define $\det(\mathbf{A}) := a_1 1$. For $n \geq 2$, we define $\det(\mathbf{A})$ recursively as

$$\det(\mathbf{A}) := \sum_{j=1}^{n} (-1)^{1+j} \mathbf{A}_{1j} \cdot \det(\widetilde{\mathbf{A}}_{1j}).$$

THe scalar $\det(\mathbf{A})$ is called the *determinant* of **A** and is also denoted by $|\mathbf{A}|$. The scalar

$$(-1)^{i+j} \det(\widetilde{\mathbf{A}}_{1j})$$

is called the cofactor of the entry of **A** in row i, column j.

- A matrix **A** is invertible iff its determinant is nonzero.

Theorem 1.15

The determinant $\det\colon \mathrm{M}_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is an alternating n-linear function. The former (alternating) means that for $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$ and any **B** obtained from **A** by interchanging any two rows of **A**,

$$\det(\mathbf{B}) = -\det(\mathbf{A}).$$

The latter (n-linearity) means that, for any scalar $k \in \mathbb{F}$ and vectors $\mathbf{u}, \mathbf{v}, \mathbf{a}_i \in \mathbb{F}^n$,

$$\det\begin{pmatrix}\mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{r-1}\mathbf{u} + k\mathbf{v} \\ \mathbf{a}_{r+1} \\ \vdots \\ \mathbf{a}_n\end{pmatrix} = \det\begin{pmatrix}\mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{r-1}\mathbf{u} \\ \mathbf{a}_{r+1} \\ \vdots \\ \mathbf{a}_n\end{pmatrix} + k\det\begin{pmatrix}\mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{r-1}\mathbf{v} \\ \mathbf{a}_{r+1} \\ \vdots \\ \mathbf{a}_n\end{pmatrix}.$$

(In fact, it can be shown that $\det\colon \mathrm{M}_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is the *unique* alternating n-linear function, such that $\det(\mathbf{I}) = 1$.)

Corollary 1.16

Let $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$. Then, for any matrix $\mathbf{B}$ obtained by adding a scalar multiple of one row/column of $\mathbf{A}$ to another, $\det(\mathbf{B}) = \det(\mathbf{A})$.

Theorem 1.17

The determinant of a square matrix can be evaluated by cofactor expansion along any row. That is, if $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$, then for any integer $1 \le i \le n$,

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} \mathbf{A}_{ij} \cdot \det(\widetilde{\mathbf{A}}_{ij}).$$

Here, $\widetilde{\mathbf{A}}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained from $\mathbf{A}$ by deleting its ith row and jth column.

Corollary 1.18

The determinant of any triangular matrix is the product of its diagonals.

Theorem 1.19

Let A be an $n \times n$ matrix. Then,

$$\det(\mathbf{A}) = \det(\mathbf{A}^\top).$$

So, the determinant of a square matrix can also be evaluated by cofactor expansion along any column.

Theorem 1.20

Let $\mathbf{A}$ be an invertible $n \times n$ matrix. Then,

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

where $\operatorname{adj}(\mathbf{A})$ is the adjugate/classical adjoint of $\mathbf{A}$. That is, the matrix whose (i,j)th entry is the (j,i)th cofactor $(-1)^{j+i}\det(\widetilde{\mathbf{A}}_{ji})$

Theorem 1.21

For any $\mathbf{A}, \mathbf{B} \in \mathrm{M}_{n\times n}(\mathbb{F})$, we have $\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$.

Definition 1.22

A linear operator T on a finite-dimensional vector space V is called *diagonalisable* iff there is an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix. A square matrix $\mathbf{A}$ is called diagonalisable iff $L_\mathbf{A}$ is diagonalisable.

Definition 1.23

Let T be a linear operator on a vector space V. A nonzero vector $\mathbf{v} \in V$ is called an *eigenvector* of T iff there exists a scalar λ such that $T(\mathbf{v}) = \lambda\mathbf{v}$. The scalar λ is called the *eigenvalue* corresponding to the eigenvector $\mathbf{v}$.
Let $\mathbf{A}$ be in $\mathrm{M}_{n\times n}(\mathbb{F})$. A nonzero vector $v \in \mathbb{F}^n$ is called an *eigenvector* of $\mathbf{A}$ iff v is an eigenvector of $L_\mathbf{A}$; that is, iff $\mathbf{A}v = \lambda v$ for some scalar λ. The scalar λ is called the eigenvalue of $\mathbf{A}$ corresponding to the eigenvector v.

Definition 1.24

Let $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$. The polynomial $f(t) = \det(\mathbf{A} - \lambda\mathbf{I}_n)$ is called the *characteristic polynomial* of $\mathbf{A}$.

- A matrix $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$ is diagonalizable iff there exists an ordered basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for $\mathbb{F}^n$ consisting of eigenvectors of $\mathbf{A}$, i.e. a eigenbasis. Furthermore, if $\mathbf{Q}$ is the $n \times n$ matrix whose jth column is $\mathbf{v}_j$, then $\mathbf{A} = \mathbf{Q}^{-1}\mathbf{D}\mathbf{Q}$ is a diagonal matrix such that d_{jj} is the eigenvalue of A corresponding to $\mathbf{v}_j$. The matrix $\mathbf{Q}$ is said to *diagonalise* $\mathbf{A}$.
- Hence, we obtain the following procedure to diagonalise a 3×3 matrix $\mathbf{A}$ with three distinct eigenvalues.
  1. Find the eigenvalues λ_1, λ_2, and λ_3 of $\mathbf{A}$. They are just the roots of the characteristic polynomial of $\mathbf{A}$. This can be done using the GC.
  2. Find an eigenvector $\mathbf{v}_j$ corresponding to each eigenvalue λ_j by finding the nullspace of $\mathbf{A} - \lambda_j \mathbf{I}$.
  3. Let $\mathbf{Q} = \left(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right)$. Then,
  $$\mathbf{D} := \mathbf{Q}^{-1} A \mathbf{Q}$$
  is a diagonal matrix.

Note

Let $\mathbf{A}$ be a 3×3 real matrix with the eigenvalue λ. Then, the cross product of two nonzero rows/columns of $\mathbf{A} - \lambda\mathbf{I}$ is an eigenvector of $\mathbf{A}$.

Theorem 1.25: The Cayley-Hamilton Theorem.

Let T be a linear operator on a finite dimensional vector space V, and let $f(t)$ be the characteristic polynomial of T. Then $f(T) = T_0$, the zero transformation. That is, T "satisfies" its characteristic equation.

Corollary 1.26: The Cayley-Hamilton Theorem for Matrices.

Let A be an $n \times n$ matrix, and let $f(t)$ be the characteristic polynomial of A. Then, $f(A) = O$, the $n \times n$ zero matrix.

G.C. Skills

Finding eigenvalues of a matrix $\mathbf{A}$ using the GC.

1. `2nd` $\Longrightarrow$ `x`$^{-1}$ (`matrix`) $\Longrightarrow$ Key in the matrix $\mathbf{A} - t\mathbf{I}$, e.g. into `[A]`.
2. Plot `Y`$_1$ = det (`[A]`).
3. `2nd` $\Longrightarrow$ `trace` $\Longrightarrow$ `2:zero` $\Longrightarrow$ Find the roots.

Chapter 2

Numerical Methods

General Information

- The parity of the degree of a real polynomial is the same as that of its number of real roots.
- Let the real polynomial p given by $p(x) = a_{2n}x^{2n} + a_{2n-1}x^{2n-1} + \cdots + a_0$ have coefficients $a_n > 0$ and $a_0 < 0$. Then, it has at least one positive and one negative root.
- To show that there a continuous function f attains a root in an interval $[a, b]$, we find two values $x < y$ in the interval (e.g. $a < b$) such that $f(a)f(b) < 0$. i.e. show that f changes sign in $[a, b]$.
- To further show that the root is *unique* in $[a, b]$, it suffices to prove that f is monotone on $[a, b]$.
- Suppose we have some function $f\colon \mathbb{R} \to \mathbb{R}$ with a root α, whose value we want to approximate. There are three ways to obtain this approximation.
  1. Linear interpolation on an interval $[a, b]$ containing α. Our approximation is
  $$\frac{a|f(b)| + b|f(a)|}{|f(a)| + |f(b)|}.$$
     - The sequence $\{x_n\}$ of approximations *always* converges to α.
     - The smaller $f''(x)$ is (i.e. the slower the gradient $f'(x)$ changes) near α, the faster the rate of convergence.
     - Error:

Concave/Gradient	Positive	Negative
Upwards $\bigcup$	underestimation	overestimation
Downwards $\bigcap$	overestimation	underestimation

Table 2.1: Approximation errors when using linear interpolation.

     - See Figure 2.1 for an illustration.

Screw trying to make nice diagonal cells. Pain. Suffering.

Note

At every iteration of linear interpolation, we must ensure that $\alpha \in [a, x_n]$. Otherwise x_n may not approximate α. If $\alpha \notin [a, x_n]$, simply consider $\alpha \in [x_n, b]$ (or any other suitable interval) instead.

Note

It is important to show which interval we are interpolating on, not just the iteratively obtained values. We can present our working using the table below.

a	$f(a)$	b	$f(b)$	$\dfrac{a\lvert f(b)\rvert + b\lvert f(a)\rvert}{\lvert f(a)\rvert + \lvert f(b)\rvert}$
a	$f(a) > 0$	b	$f(b) < 0$	x_1
x_1	$f(x_1) > 0$	b	$f(b) < 0$	x_2
x_1	$f(x_1) > 0$	x_2	$f(x_2) < 0$	x_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 2.2: Required working for linear interpolation.

1. Fixed-point Iteration. First select a function $F\colon \mathbb{R} \to \mathbb{R}$, such that $F(\alpha) = \alpha$, and choose some initial approximation x_0 to α. Then, we recursively define $x_{n+1} \coloneqq F(x_n)$. We want $x_n \to \alpha$.
   - Convergence behavior

Behaivor of $\lvert F'(x)\rvert$	Converges?	Rate of convergence
$\lvert F'(x)\rvert < 1$ and is small near α	✓	fast
$\lvert F'(x)\rvert < 1$ but is close to 1 near α	✓	slow
$\lvert F'(x)\rvert \geq 1$ near α	×	-

Table 2.3: Convergence behavior of fixed-point iterations.

   - See Figure 2.2 for an illustration.

Note

We must write out *all* iterations, not just the final two. The working below is sufficient.

Let $x_0 =$ ___ and $x_{n+1} = F(x_n)$, $x \geq 0$.

$$\begin{aligned} x_1 &= \_\_\_ \\ x_2 &= \_\_\_ \\ &\vdots \\ x_{m-1} &= \_\_\_ \\ x_m &= \_\_\_ \end{aligned}$$

Therefore, $\alpha = x_m$ (k d.p.).

1. The Newton-Raphson Method. Let α be a root of the function $f\colon \mathbb{R} \to \mathbb{R}$. The Newton-Raphson formula is
$$x_{n+1} \coloneqq x_n - \frac{f(x_n)}{f'(x_n)}.$$
   - The Newton-Raphson method fails in the following cases.
     (a) The gradient at x_0 is too gentle.
     (b) The gradient changes too rapidly.
     (c) The initial approximation x_0 is too far from the root α.
     (d) There is a turning point between the initial approximation x_0 and the root α.

(e) There is a point of inflection — where the concavity changes/the sign of $f''(x)$ changes.

– Error:

Concave/Gradient	Positive	Negative
Upwards $\bigcup$	overestimation	underestimation
Downwards $\bigcap$	underestimation	overestimation

Table 2.4: Approximation errors when using the Newton-Raphson method.

– See Figure 2.3 for an illustration.

Note

We must write out *all* iterations, not just the final two. One way to present our working is as follows.

Let $x_0 = \_\_$ and $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = \_\_$, $x \geq 0$.

$$\begin{aligned} x_1 &= \_\_ \\ x_2 &= \_\_ \\ &\vdots \\ x_{m-1} &= \_\_ \\ x_m &= \_\_ \end{aligned}$$

Therefore, $\alpha = x_m$ (k d.p.).

Note

Suppose a question asks for the approximation of a root to k significant figures/k decimal places. Then:

1. We leave our iterative approximations x_n to at least $k+2$ significant figures/$k+2$ decimal places.

2. We continue the iterative process till two consecutive ones agree up to k significant figures/k decimal places.

G.C. Skills

Linear interpolation: finding an approximation to a root in $[a, b]$ up to n decimal places.

1. $Y_1 = f(x)$,
2. $a \to A$ and $b \to B$,
3. $\dfrac{B|Y_1(A)| + A|Y_1(B)|}{|Y_1(A)| + |Y_1(B)|}$,
4. Ans $\to A$ or B (choose the one that has the opposite sign to Ans),
5. Repeat steps 4 to 5,
6. Terminate this process when the approximations are consistent up to n decimal places.

You can freely enter any function and shift the initial values in the Desmos graphs below!

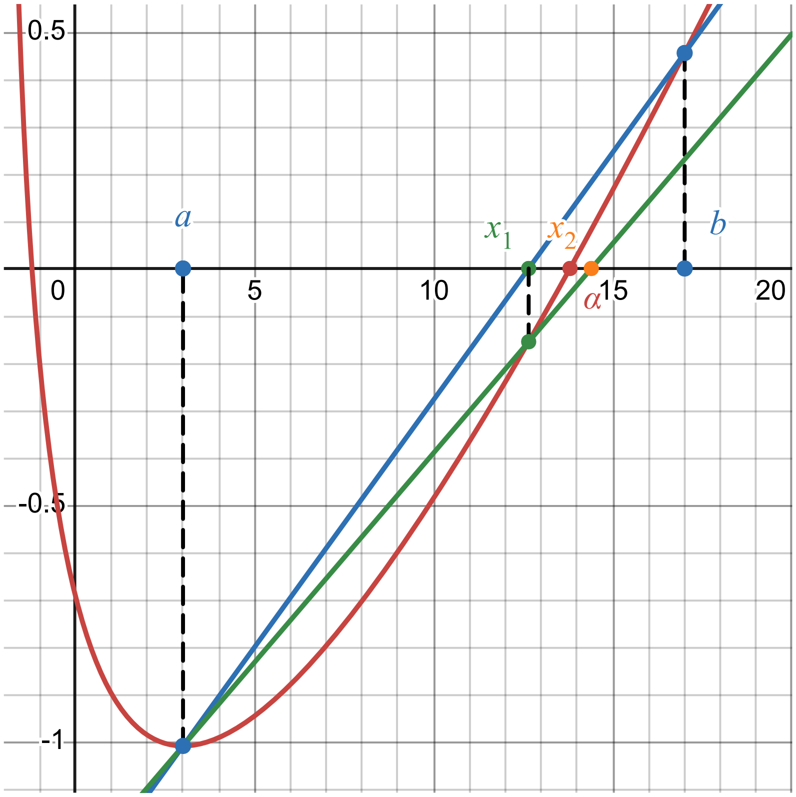

Figure 2.1: An illustration of linear interpolation (Desmos).

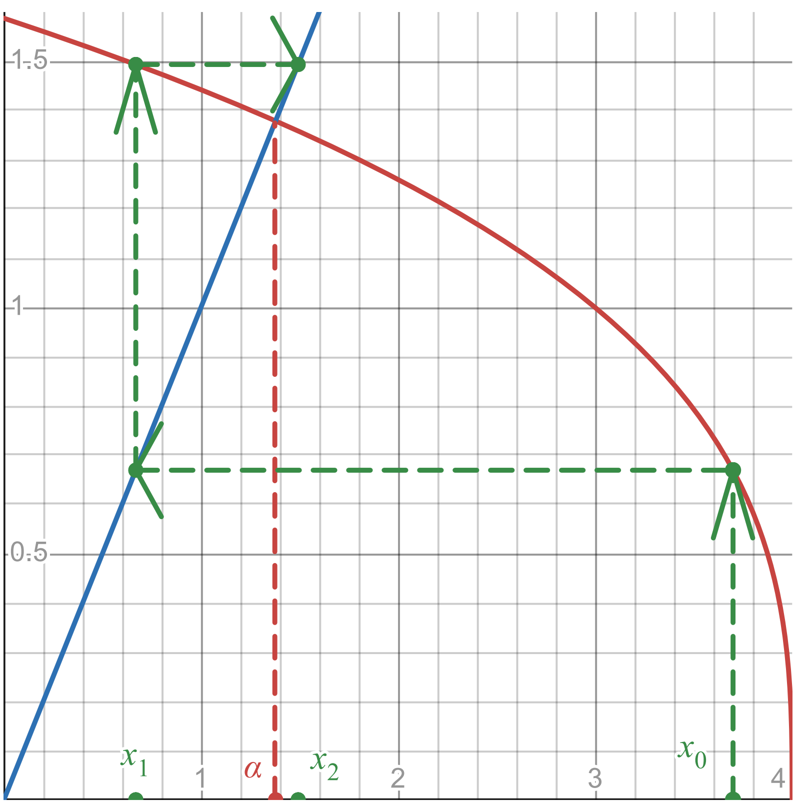

Figure 2.2: An illustration of fixed-point iteration (Desmos).

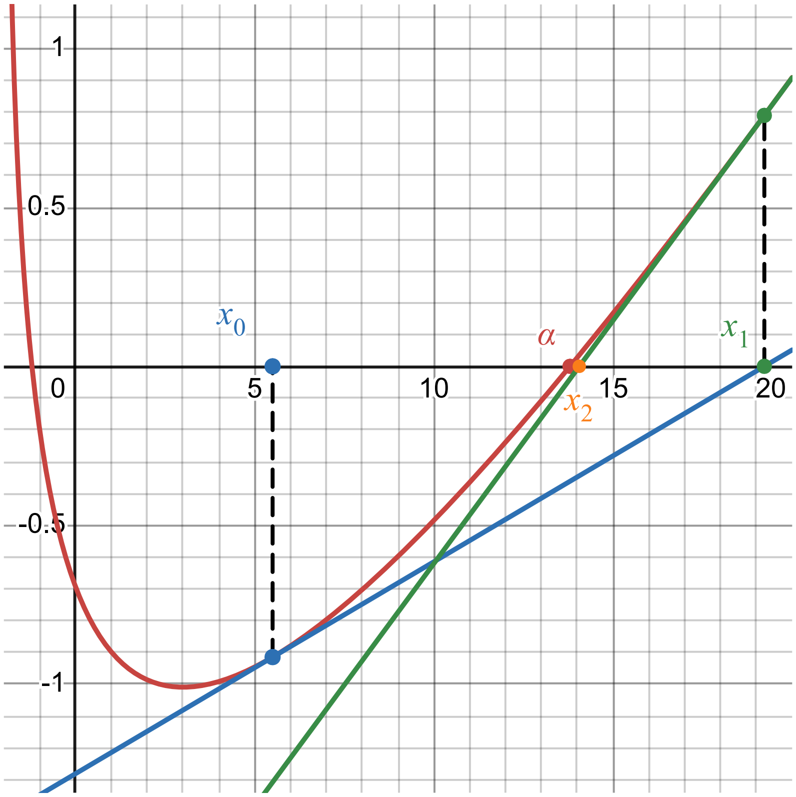

Figure 2.3: An illustration of Newton's Method (Desmos).