Chapter 1

Linear Algebra

1.1 Vector spaces, subspaces, linear combinations, bases, linear transformations

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 $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ in S and scalars $a_1, a_2, \ldots, a_n$ in $\mathbb{F}$ such that

$$v = \sum_{i=1}^{n} a_i \mathbf{u}_i.$$

In this case we also say that v is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ 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 $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$in S and scalars $a_1, a_2, \ldots, a_n$ not all zero, such that

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + a_n\mathbf{u}_n = \mathbf{0}.$$

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.

Definition 1.7

Let V and W be vector spaces. We call a function $T\colon V \to W$ a *linear transformation from V to W* iff $T(c\mathbf{x} + \mathbf{y}) = cT(\mathbf{x}) + T(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in V$ and $c \in \mathbb{F}$.

Theorem 1.8: The Rank-Nullity Theorem.

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

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

1.2 Matrices and systems of linear equations

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

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

Definition 1.9

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

Theorem 1.10

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

Theorem 1.11

A system $\mathbf{Ax} = \mathbf{0}$ of m linear equations in n unknowns has a solution space of dimension $n - \text{rank}(A)$.

Definition 1.12

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

Theorem 1.13: The Rouché-Capelli Theorem.

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

Definition 1.14

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.

General Information

- 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 reduces a matrix to its rref form.
- Gaussian elimination always reduces $(\mathbf{A} \,|\, \mathbf{I}_n) \to (\mathbf{I}_n \,|\, \mathbf{A}^{-1})$.
- Let $\mathbf{A} \coloneqq \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$ be $m \times n$ matrix, and $\mathbf{A}' \coloneqq \begin{pmatrix} \mathbf{a}'_1 & \mathbf{a}'_2 & \cdots & \mathbf{a}'_n \end{pmatrix}$ its rref. Then, $\{\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \ldots, \mathbf{a}_{k_m}\}$ is linearly independent iff $\{\mathbf{a}'_{k_1}, \mathbf{a}'_{k_2}, \ldots, \mathbf{a}'_{k_m}\}$ is. Moreover, the row space of $\mathbf{A}$ and $\mathbf{A}'$ are clearly identical.
- (Example) To find a basis for the intersection of the column spaces of $\mathbf{A}, \mathbf{B} \in \mathrm{M}_{n\times n}(\mathbb{F})$, we 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 columns of $\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$ and $\begin{pmatrix} \mathbf{A}' & \mathbf{B}' \end{pmatrix}$, respectively. We compare the columns of $\mathbf{A}'$ and $\mathbf{B}'$ to find a basis $\beta' \coloneqq \{\mathbf{c}'_{i_1}, \mathbf{c}'_{i_2}, \ldots, \mathbf{c}'_{i_r}\}$ for the intersection of the column spaces of $\mathbf{A}'$ and $\mathbf{B}'$. Then, $\beta \coloneqq \{\mathbf{c}_{i_1}, \mathbf{c}_{i_2}, \ldots, \mathbf{c}_{i_r}\}$ is a basis for the intersection of the column spaces of $\mathbf{A}$ and $\mathbf{B}$.

1.3 Determinants

Definition 1.15

Let $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$. If $n = 1$, so that $A = (a_{11})$, we define $\det(\mathbf{A}) \coloneqq a_{11}$. For $n \geq 2$, we define $\det(\mathbf{A})$ recursively as

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

The scalar $\det(\mathbf{A})$ is called the *determinant* of $\mathbf{A}$ 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 $\mathbf{A}$ in row i, column j.

Note

A matrix $\mathbf{A}$ is invertible iff its determinant is nonzero.

Theorem 1.16

The determinant $\det\colon \mathrm{M}_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is an alternating n-linear function.

(a) Alternating: For $\mathbf{A} \in \mathrm{M}_{n\times n}(\mathbb{F})$ and any $\mathbf{B}$ obtained from $\mathbf{A}$ by interchanging any two rows of $\mathbf{A}$,

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

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

Corollary 1.17

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 1.18

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 $1 \leq i \leq n$,

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

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 1.19
The determinant of any triangular matrix is the product of its diagonals.

Theorem 1.20
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 1.21
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 1.22
For any $\mathbf{A}, \mathbf{B} \in \mathrm{M}_{n\times n}(\mathbb{F})$, we have $\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$.

1.4 Diagonalisation

Definition 1.23
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 1.24
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 1.25
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 $\mathbf{A}$ — 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 reducing $\mathbf{A} - \lambda_j\mathbf{I}$.

3. Let $\mathbf{Q} = \begin{pmatrix}\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\end{pmatrix}$. Then,
$$\mathbf{D} \coloneqq \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$$
is a diagonal matrix.

Note

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

Theorem 1.26: The Cayley-Hamiliton 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 1.27: The Cayley-Hamiliton Theorem for Matrices.

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

G.C. Skills

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

1. `2nd` $\Longrightarrow$ $\mathtt{x}^{-1}$ (`matrix`) $\Longrightarrow$ Key in the matrices $\mathbf{A}$ and $\mathbf{I}_3$ into `[A]` and `[I]`, respectively.
2. Plot $\mathtt{Y}_1 = \det(\mathtt{[A]})$.
3. `2nd` $\Longrightarrow$ `trace` $\Longrightarrow$ `2:zero` $\Longrightarrow$ Find the roots.

1.5 Miscellaneous

An asterisk denotes the conjugate transpose.

Theorem 1.28

Let $\mathbf{M} \in \mathrm{M}_{n\times n}(\mathbb{K})$ be Hermitian (i.e. $\mathbf{M}^* = \mathbf{M}$), with eigenvectors $\mathbf{u}$ and $\mathbf{v}$ that correspond to the eigenvalues λ and μ. Then, $\mathbf{u}$ and $\mathbf{v}$ are orthogonal with respect to the standard inner product, if $\lambda \neq \mu^*$.

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be eigenvectors of $\mathbf{M}$. Then,
$$\langle \mathbf{u}, \mu\mathbf{v}\rangle = (\mathbf{M}\mathbf{v})^*\mathbf{u} = (\mathbf{v}^*\mathbf{M}^*)\mathbf{u} = \mathbf{v}^*(\mathbf{M}^*\mathbf{u}) = \mathbf{v}^*(\lambda\mathbf{u}) = \langle \lambda\mathbf{u}, \mathbf{v}\rangle.$$
As such, $(\lambda - \mu^*)\langle \mathbf{u}, \mathbf{v}\rangle = 0$. Hence, $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. □

Example 1.1

Consider a computer that rounds each calculated value to n decimal places, which is then used in later calculations as if it were exact. Perform, for $n = 3$ and $n = 4$, this procedure to find the solution $\mathbf{x} = \begin{pmatrix}x_1 & x_2 & x_3\end{pmatrix}^\top$ to $\mathbf{A}\mathbf{x} = \mathbf{b}$. Then, find $\sum_{i=1}^{3} \delta_i^2$ where δ_i is the difference between the exact value of x_i and the one found by the computer: this gives a measure for the accuracy of the calculated values. Comment on the difference in results.

One extra decimal place of accuracy in the working (a factor of 10) had led to a significant increase in the measure of accuracy (by a factor of around 250).

1.6 Conics

Given a conic section defined by $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$, we may wish to find its lines of symmetry, center, radii, etc. The core idea is simple: complete the square, to express $Ax^2 + Bxy + Cy^2$ in the form $a(x')^2 + b(y')^2$, for some linear combinations x' and y' of x and y. Then, the initial equation becomes $a(x')^2 + b(y')^2 + d(x') + e(y') + F = 0$, which is easily reduced to the standard form for conics. Before that, we need to develop some machinery.

Definition 1.29

Let $\mathbb{F}$ be a field not of characteristic two. A function $K\colon \mathbb{F}^n \to \mathbb{F}$ is called a *quadratic form on* $\mathbb{F}^n$ if there exists a symmetric matrix $A \in \mathrm{M}_{n\times n}(\mathbb{F})$, such that $K(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$.

Note

Let $\mathbb{F}$ be a field not of characteristic two and scalars $a_i \in \mathbb{F}$. The polynomial $f\colon \mathbb{F}^n \to \mathbb{F}$ given by $f(t_1, t_2, \ldots, t_n) = \sum_{i\le j} a_{ij} t_i t_j$ is a quadratic form. In fact, the matrix $\mathbf{A}$ with

$$\mathbf{A}_{ij} = \begin{cases} a_{ii} & \text{if } i = j \\ a_{ij}/2 & \text{if } i \neq j \end{cases}$$

gives us our desired quadratic form $K(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$.

Theorem 1.30

Let $K(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$ be a quadratic form on a finite-dimensional real inner product space V. There exists an orthonormal eigenbasis $\beta \coloneqq \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for $\mathbf{A}$ and eigenvalues λ_i, such that $K\left(\sum_{i=1}^n t_i \mathbf{v}_i\right) = \sum_{i=1}^n \lambda_i t_i^2$ for all $t_i \in \mathbb{R}$.

We now return to our initial problem on conics. We first diagonalise the matrix

$$\begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}$$

to find its eigenvalues λ and μ, then the corresponding unit eigenvectors $\mathbf{u} = \begin{pmatrix}\alpha & \beta\end{pmatrix}^\top$ and $\mathbf{v} = \begin{pmatrix}\gamma & \delta\end{pmatrix}^\top$. Then,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix}\mathbf{u} & \mathbf{v}\end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} \alpha t_1 + \gamma t_2 \\ \beta t_1 + \delta t_2 \end{pmatrix}.$$

Furthermore, $Ax^2 + Bxy + Cy^2 = \lambda t_1^2 + \mu t_2^2$ by 1.28. Therefore,

$$\lambda t_1^2 + \mu t_2^2 + D(\alpha t_1 + \gamma t_2) + E(\beta t_1 + \delta t_2) + F = 0$$

$$\lambda\left(t_1 + \frac{D\alpha + E\beta}{2\lambda}\right)^2 + \mu\left(t_2 + \frac{D\gamma + E\delta}{2\mu}\right)^2 - \frac{(D\alpha + E\beta)^2}{4\lambda} - \frac{(D\gamma + E\delta)^2}{4\mu} + F = 0 \qquad (\text{if } \lambda, \mu \neq 0)$$

gives an equivalent form for our conic section.

Since our basis $\{\mathbf{u}, \mathbf{v}\}$ is orthonormal, the above change of basis from the standard ordered basis to $\{\mathbf{u}, \mathbf{v}\}$ is an isometry (that maps $\mathbf{0}$ to itself). It is obtained via a rotation and/or reflection. Hence, the radii are $\sqrt{\lambda/k}$ and $\sqrt{\mu/k}$, for $4k = (D\alpha + E\beta)^2/\lambda + (D\gamma + E\delta)^2/\mu - F$; the center is $\left(-\frac{D\alpha+E\beta}{2\lambda}, -\frac{D\gamma+E\delta}{2\mu}\right)$. Notice that $t_1 = 0$ and $t_2 = 0$ are parallel to the axes of symmetry. i.e. $\mathbf{u}$ and $\mathbf{v}$ are parallel to the axes of symmetry of our conic. As such, the lines of symmetry are

$$y + \frac{D\gamma + E\delta}{2\mu} = \frac{\beta}{\alpha}\left(x + \frac{D\alpha + E\beta}{2\lambda}\right) \qquad \text{and} \qquad y + \frac{D\gamma + E\delta}{2\mu} = \frac{\delta}{\gamma}\left(x + \frac{D\alpha + E\beta}{2\lambda}\right).$$

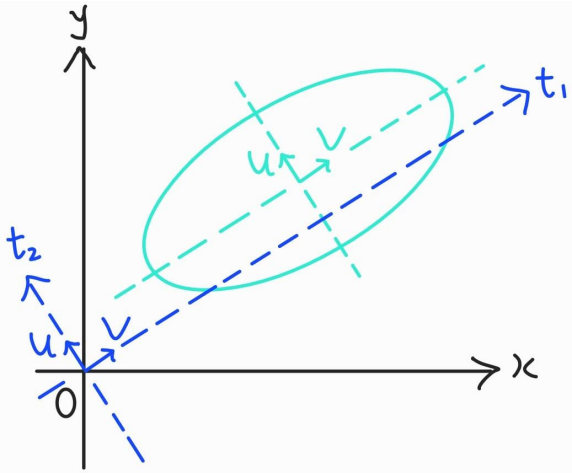

Figure 1.1: Rotated and translated conic.

Note

Without orthonormality, i.e. an isometric change of coordinates, even if we were successful in reducing our conic to the form $a(x')^2 + b(y')^2 = f$, it may not prove to be a useful form. Consider the ellipse $2x^2 + 2xy + y^2 = 1$. Clearly, $x^2 + (x + y)^2 = 1$. But, this gives a circle in $(x, x + y)$ coordinates; we can't deduce much about our initial conic. So, little meaning is found in such a factorisation.

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 $[a, b]$. Then, *by continuity*, a root of f must lie in $[a, b]$.
- To further show that the root is *unique* in $[a, b]$, it suffices to prove that f is *strictly* 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 2.2 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	$f(b)$	$\frac{a\lvert f(b)\rvert + b\lvert f(a)\rvert}{\lvert f(a)\rvert + \lvert f(b)\rvert}$
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.

2. 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

Behvaior 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 2.3 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.), since $f(x_m - 0.0\cdots05)f(x_m + 0.0\cdots05) = \_\_\_ < 0$.

3. 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 2.4 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.), since $f(x_m - 0.0\cdots05)f(x_m + 0.0\cdots05) = \_\_\_ < 0$.

Note

Explain whether $x_0 = \_\_\_$ is a suitable starting value for using the Newton-Raphson method to find an approximation to α.

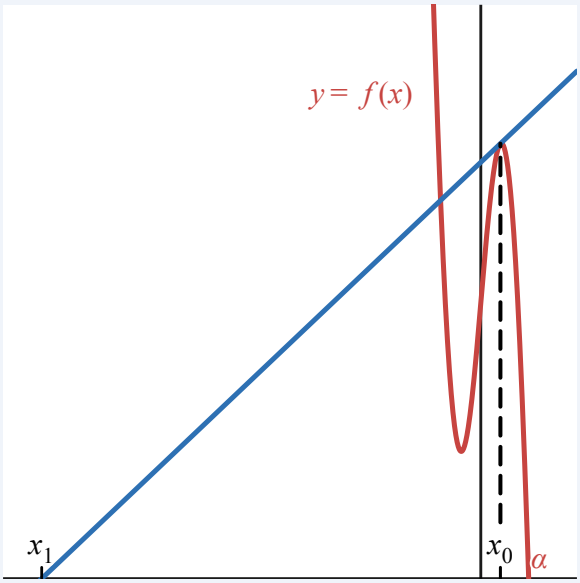

Figure 2.1: (Desmos)

1. Since $x_0 =$ ___ is very close to the stationary point, the tangent to the curve $y = f(x)$ has a very gentle gradient. Thus, it cuts the x-axis far away from the initial approximation.
2. Furthermore, as $x_0 =$ ___ is to the left/right of the minimum/maximum point $x =$ ___, the values of the gradient $f'(x_n)$ will be negative/positive for all $n \geq 0$. Hence, x_n converges to the root β instead α.

(The second point may be omitted if it is irrelevant.)

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.

Note

Perform _____ (e.g. linear interpolation) to obtain an approximation for α, correct to two decimal places. Justify whether this approximation is sufficiently accurate.

Suppose our approximation is some $a = 1.00$, then we note the sign of f at $a \pm 0.005$. (For an arbitrary number of s.f. or d.p., simply adjust the value 0.005 accordingly. E.g. for 3 d.p. we instead use 0.0005). Our working should look similar to the following:

> Since $f(0.995) =$ ___ < 0 and $f(1.005) =$ ___ > 0, we conclude that 1.00 is a sufficiently accurate approximation, at 2 d.p..

Note

The *error obtained* when using an approximation should be the *absolute* difference of the true value and the approximation.

Note

Use the results of part (i) and the differential equation $dy/dx = \sin(xy)$ to estimate the x-coordinate x_P of P.

The maximum point occurs when $dy/dx = \sin(xy) = 0$, i.e. $xy = k\pi$ where $k \in \mathbb{Z}$. From (i), $\frac{dy}{dx}\big|_{x=5/3} \approx 0.643 > 0$ so $y_P > y(5/3) \approx 2.0468$. Hence, $x_P \approx \pi/2.04679402 = 1.53$ (3.s.f).

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!

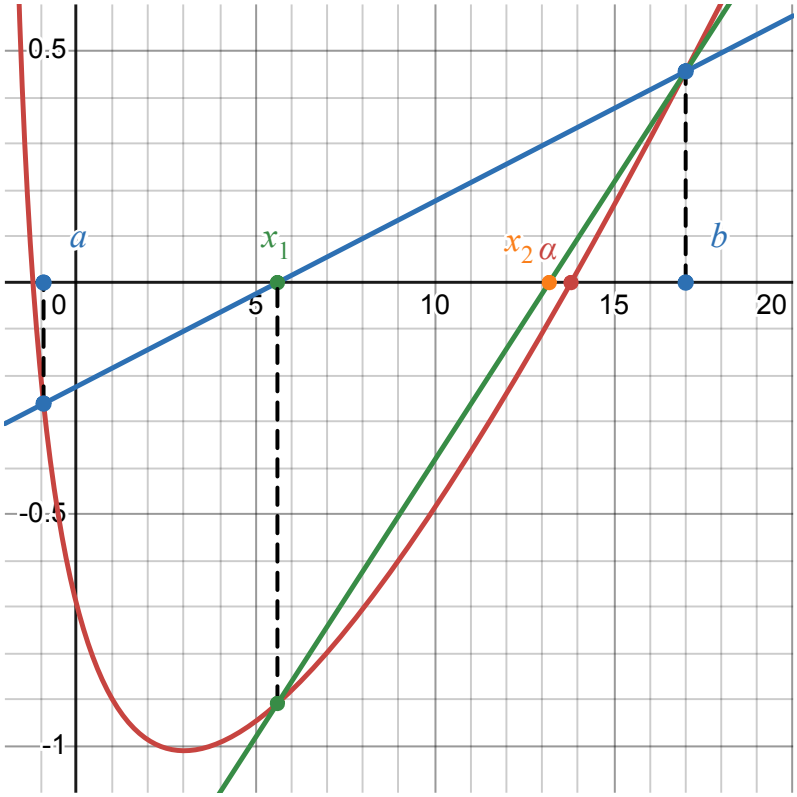

Figure 2.2: An illustration of linear interpolation (Desmos).

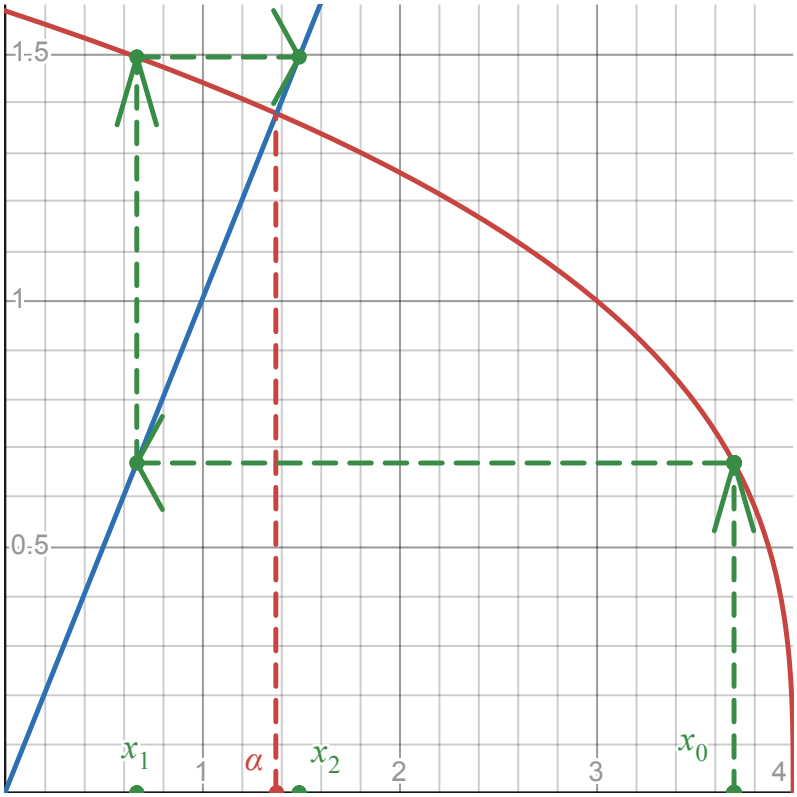

Figure 2.3: An illustration of fixed-point iteration (Desmos).

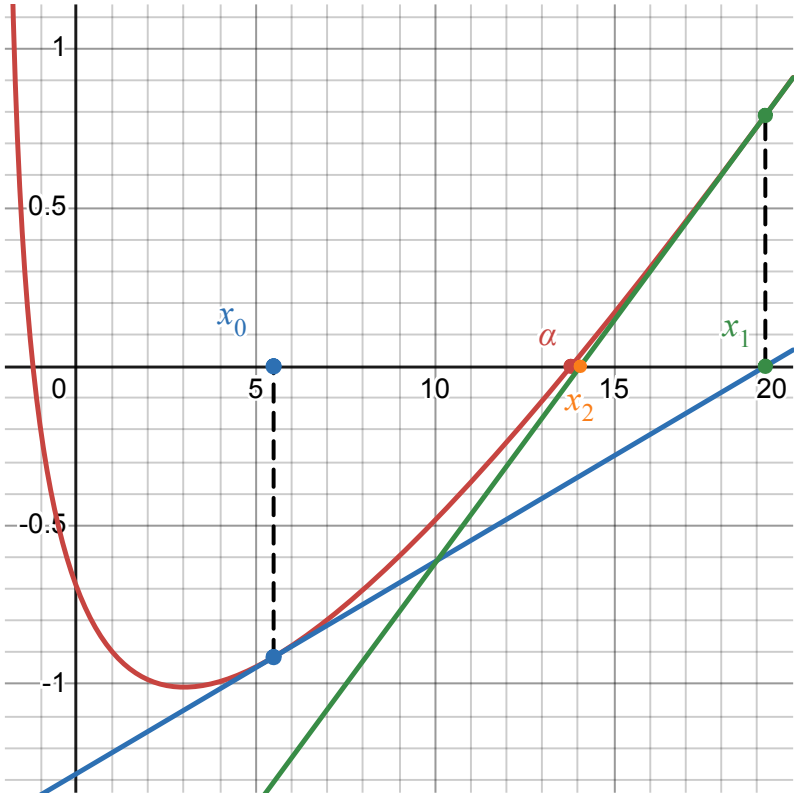

Figure 2.4: An illustration of Newton's Method (Desmos).