# Vectors & Matrices - Definitions

Lectured by Prof Paul Linden

Michaelmas 2013

## 1 Complex Numbers

**Definition 1** (Complex Logarithm)**.** For $z \in \mathbb{C}, \log z = \log |z| + i \arg z$.

**Definition 2** (Complex Powers)**.** For $z \neq 0$ and $z, w \in \mathbb{C}$, define $z$ raised to power $w$ as $z^w = \exp(w \log z)$.

## 2 Vector Algebra

**Definition 3** (Vector)**.** A *vector* is a quantity specified by a positive magnitude and a direction in space.

**Definition 4** (Vector field)**.** $\mathbf{F}(\mathbf{x})$ is a vector function of space. e.g. an electric field.

**Definition 5** (Vector space)**.** A *vector space* over the real numbers is the set of elements $V$ under two binary operations, and with the null, inverse and identity element satisfying:

(i) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (commutative)

(ii) $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$ (associative)

(iii) $\mathbf{a} + \mathbf{0} = \mathbf{a}$ (null vector)

(iv) $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$ (additive inverse vector)

(v) $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$ (scalar multiplication distributive over scalar addition)

(vi) $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$ (scalar multiplication distributive over vector addition)

(vii) $\lambda(\mu\mathbf{a}) = (\lambda\mu)\mathbf{a}$ (scalar multiplication 'associative')

(viii) $1\mathbf{a} = \mathbf{a}$ (multiplicative identity)

**Definition 6** (Scalar product)**.** The *scalar product* or *dot product* of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined geometrically to be the real (scalar) number:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$$

where $0 \leq \theta \leq \pi$ is the non-reflex angle between $\mathbf{a}$ and $\mathbf{b}$ once they have been placed 'tail to tail'.

**Definition 7** (Vector product)**.** The *vector product* or *cross product* of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined geometrically to be a vector such that:

(i)
$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \theta$$
with $0 \le \theta \le \pi$ defined as with the scalar product;

(ii) $\mathbf{a} \times \mathbf{b}$ is perpendicular/orthogonal to both $\mathbf{a}$ and $\mathbf{b}$;

(iii) $\mathbf{a} \times \mathbf{b}$ has the sense/direction defined by the 'right-hand rule', i.e. point the index finger in the direction of $\mathbf{a}$, the middle finger in the direction of $\mathbf{b}$, then $\mathbf{a} \times \mathbf{b}$ is in the direction of the thumb.

**Definition 8** (Triple products)**.** Given three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, we can form two triple products.
Scalar triple product
$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$
Vector triple product
$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$$

**Definition 9** (Linear independence)**.** A set of $m$ vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_m\}, \mathbf{v}_j \in \mathbb{R}^n, j = 1, 2, \dots m$ is *linearly independent* if for all scalars $\lambda_j \in \mathbb{R}, j = 1, 2, \dots m$

$$\sum_{i=1}^{m} \lambda_i \mathbf{v}_i = \mathbf{0} \implies \lambda_i = 0 \quad \text{for} \quad i = 1, 2, \dots m$$

Otherwise the vectors are said to be *linearly dependent.*

**Definition 10** (Spanning set)**.** A subset $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots \mathbf{u}_m\}$ of vectors in $\mathbb{R}^n$ is a *spanning set* for $\mathbb{R}^n$ if for every vector $\mathbf{v} \in \mathbb{R}^n$, there exist scalars $\lambda_j \in \mathbb{R}, j = 1, 2, \dots m$, such that
$$\mathbf{v} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \ldots + \lambda_m \mathbf{u}_m$$

**Definition 11** (Basis)**.** A linearly independent subset of vectors that spans $\mathbb{R}^n$ is a *basis* for $\mathbb{R}^n$.

**Definition 12** (Components)**.** Given a basis $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots \mathbf{u}_m\}$ of $\mathbb{R}^n$, for every vector $\mathbf{v} \in \mathbb{R}^n$ there exist *unique* scalars $\lambda_j \in \mathbb{R}, j = 1, 2, \dots n$ such that

$$\mathbf{v} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \ldots + \lambda_n \mathbf{u}_n$$

The $\lambda_j$ are the *components* of $\mathbf{v}$ with respect to the ordered basis $S$.

**Definition 13** (Dimension)**.** The *dimension* of a space is the number of vectors in a basis of the space.

**Definition 14** (Scalar product for $\mathbb{R}^n$)**.** For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define the *scalar product*

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

**Definition 15** (Euclidean norm)**.** The *length* or *Euclidean norm* of a vector $\mathbf{x} \in \mathbb{R}^n$ is

$$|\mathbf{x}| \equiv (\mathbf{x} \cdot \mathbf{x})^{\frac{1}{2}} = \left( \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$$

**Definition 16** (Interior angle)**.** The *interior angle* $\theta$ between two vectors $\mathbf{x}$ and $\mathbf{y}$ is

$$\theta = \arccos \left( \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|} \right)$$

**Definition 17** (Scalar product for $\mathbb{C}^n$)**.** For $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$, define the *scalar product*

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i^* v_i = u_1^* v_1 + u_2^* v_2 + \ldots + u_n^* v_n$$

**Definition 18** (Orthogonal)**.** Non-zero vectors $\mathbf{x}, \mathbf{y}$ are *orthogonal* if $\mathbf{x} \cdot \mathbf{y} = 0$.

**Definition 19** (Suffix notation)**.** Using suffices, we can abbreviate expressions involving vectors by following these rules:

- if a suffix appears once it is taken to be a free suffix and ranged through,
- if a suffix appears twice it is taken to be a dummy suffix and summed over,
- if a suffix appears more than twice *in one term*, something has gone wrong.

**Definition 20** (Kronecker delta)**.** The *Kronecker delta* $\delta_{ij}, i, j = 1, 2, 3$ is defined by

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

**Definition 21** (Levi-Civita symbol)**.** The *Levi-Civita symbol* or *alternating tensor* $\epsilon_{ijk}$ is defined by

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } i, j, k \text{ is an even permutation of } 1\ 2\ 3; \\ -1 & \text{if } i, j, k \text{ is an odd permutation of } 1\ 2\ 3; \\ 0 & \text{otherwise} \end{cases}$$

**Definition 22** (Subspace)**.** A non-empty subset $U$ of the elements of a vector space $V$ is called a *subspace* of $V$ if $U$ is a vector space under the same operations as are used to define $V$.

# 3 Matrices and Linear Maps

**Definition 23** (Map)**.** Let $A, B$ be sets. A *map* $\mathcal{T}$ of $A$ into $B$ is a rule that assigns to each $x \in A$ a unique $x' \in B$. We write

$$\mathcal{T} : A \to B \text{ and/or } x \mapsto x' = \mathcal{T}(x)$$

$A$ is the *domain* of $\mathcal{T}$.
$B$ is the *range*, or *codomain* of $\mathcal{T}$.
$\mathcal{T}(x) = x'$ is the *image* of $x$ under $\mathcal{T}$.
$\mathcal{T}(A)$ is the *image* of $A$ *under* $\mathcal{T}$, i.e. the set of all image points $x' \in B$ of $x \in A$.

**Definition 24** (Linear map)**.** Let $V, W$ be real vector spaces, e.g. $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. The map $\mathcal{T} : V \to W$ is a *linear map* or *linear transformation* if for all $\mathbf{a}, \mathbf{b} \in V$ and $\lambda, \mu \in \mathbb{R}$,

(i) $\mathcal{T}(\mathbf{a} + \mathbf{b}) = \mathcal{T}(\mathbf{a}) + \mathcal{T}(\mathbf{b})$

(ii) $\mathcal{T}(\lambda\mathbf{a}) = \lambda\mathcal{T}(\mathbf{a})$

or equivalently if $\mathcal{T}(\lambda\mathbf{a} + \mu\mathbf{b}) = \lambda\mathcal{T}(\mathbf{a}) + \mu\mathcal{T}(\mathbf{b})$

**Definition 25** (Rank)**.** The *rank* of a linear map $\mathcal{T} : V \to W$ is the dimension of the image

$$r(\mathcal{T}) = \mathrm{dim}\mathcal{T}(V)$$

**Definition 26** (Kernel)**.** For a linear map $\mathcal{T} : V \to W$, the subset of $V$ that maps to the zero element in $W$ is called the *kernel* or *null space* of $\mathcal{T}$

$$K(\mathcal{T}) = \{\mathbf{v} \in V : \mathcal{T}(\mathbf{v}) = \mathbf{0} \in W\}$$

**Definition 27** (Nullity)**.** The *nullity* of a linear map $\mathcal{T} : V \to W$ is the dimension of the kernel

$$n(\mathcal{T}) = \mathrm{dim}K(\mathcal{T})$$

**Definition 28** (Composite map)**.** Suppose that $\mathcal{S} : U \to V$ and $\mathcal{T} : V \to W$ are linear maps. The *composite* or *product* map $\mathcal{TS}$ is the map $\mathcal{TS} : U \to W$ such that

$$\mathbf{u} \mapsto \mathbf{w} = \mathcal{T}(\mathcal{S}(\mathbf{u}))$$

**Definition 29** (Transpose)**.** If $\mathrm{A} = \{A_{ij}\}$ is a $m \times n$ matrix, then its *transpose* $\mathrm{A}^\mathrm{T}$ is defined to be a $n \times m$ matrix with elements

$$(\mathrm{A}^\mathrm{T})_{ij} = (\mathrm{A})_{ji} = \mathrm{A}_{ji}$$

**Definition 30** (Hermitian conjugate)**.** The *Hermitian conjugate* or *conjugate transpose* or *adjoint* of a matrix $\mathrm{A} = \{A_{ij}\}$, where $A_{ij} \in \mathbb{C}$ is

$$\mathrm{A}^\dagger = (\mathrm{A}^\mathrm{T})^* = (\mathrm{A}^*)^\mathrm{T}$$

**Definition 31** (Symmetric)**.** A square $n \times n$ matrix $\mathrm{A} = \{A_{ij}\}$ is *symmetric* if

$$\mathrm{A} = \mathrm{A}^\mathrm{T}$$

**Definition 32** (Hermitian)**.** A square $n \times n$ (complex) matrix $\mathrm{A} = \{A_{ij}\}$ is *Hermitian* if

$$\mathrm{A} = \mathrm{A}^\dagger$$

**Definition 33** (Antisymmetric)**.** A square $n \times n$ matrix $\mathrm{A} = \{A_{ij}\}$ is *antisymmetric* if

$$\mathrm{A} = -\mathrm{A}^\mathrm{T}$$

**Definition 34** (Skew-Hermitian)**.** A square $n \times n$ (complex) matrix $\mathrm{A} = \{A_{ij}\}$ is *skew-Hermitian* if

$$\mathrm{A} = -\mathrm{A}^\dagger$$

**Definition 35** (Trace)**.** The *trace* of a square $n \times n$ matrix $\mathrm{A} = \{A_{ij}\}$ is equal to the sum of the diagonal elements

$$Tr(\mathrm{A}) = A_{ii}$$

**Definition 36** (Unit matrix)**.** The *unit* or *identity* $n \times n$ matrix is

$$\mathrm{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

**Definition 37** (Inverse)**.** Let A be a $m \times n$ matrix. A $n \times m$ matrix B is a *left inverse* of A if BA = I. A $n \times m$ matrix C is a *right inverse* of A if AC = I.

**Definition 38** (Invertible)**.** An $n \times n$ matrix A is *invertible* if there exists an $n \times n$ matrix B such that

$$\mathrm{BA} = \mathrm{AB} = \mathrm{I}$$

The matrix B is called the *inverse* of A and is denoted $\mathrm{A}^{-1}$.

**Definition 39** (Orthogonal)**.** An $n \times n$ real matrix $\mathrm{A} = \{A_{ij}\}$ is *orthogonal* if

$$\mathrm{A}\mathrm{A}^{\mathrm{T}} = \mathrm{I} = \mathrm{A}^{\mathrm{T}}\mathrm{A}$$

**Definition 40** (Unitary)**.** A complex square matrix U is *unitary* if its Hermitian conjugate is equal to its inverse

$$\mathrm{U}^{\dagger} = \mathrm{U}^{-1}$$

**Definition 41** (Determinant)**.** The *determinant* of a $3 \times 3$ matrix A is given by

$$\det \mathrm{A} \equiv |\mathrm{A}| = \epsilon_{ijk} A_{i1} A_{j2} A_{k3}$$

**Definition 42** (Levi-Civita symbol - generalised)**.**

$$\epsilon_{j_1 j_2 \dots j_n} = \begin{cases} 1 & \text{if } (j_1, j_2, \dots j_n) \text{ is an even permutation of } (1, 2, \dots n); \\ -1 & \text{if } (j_1, j_2, \dots j_n) \text{ is an odd permutation of } (1, 2, \dots n); \\ 0 & \text{otherwise (if any two labels are the same)} \end{cases}$$

**Definition 43** (Determinant - generalised)**.**

$$\det \mathrm{A} = \epsilon_{i_1 i_2 \dots i_n} A_{i_1 1} A_{i_2 2} \dots A_{i_n n} = \sum_{\sigma \in S_n} \epsilon(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \dots A_{\sigma(n)n}$$

where the $i_k$ are summed over (summation convention).

**Definition 44.** For a square $n \times n$ matrix $\mathrm{A} = \{A_{ij}\}$, define $\mathrm{A}^{ij}$ to be the $(n-1) \times (n-1)$ square matrix obtained by eliminating the $i$th row and the $j$th column of A.

$$\mathrm{A}^{ij} = \begin{pmatrix} A_{11} & \dots & A_{1(j-1)} & A_{1(j+1)} & \dots & A_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{(i-1)1} & \dots & A_{(i-1)(j-1)} & A_{(i-1)(j+1)} & \dots & A_{(i-1)n} \\ A_{(i+1)1} & \dots & A_{(i+1)(j-1)} & A_{(i+1)(j+1)} & \dots & A_{(i+1)n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{n(j-1)} & A_{n(j+1)} & \dots & A_{nn} \end{pmatrix}$$

**Definition 45** (Minor)**.** The *minor* $M_{ij}$ of the $ij$th element of a square matrix A is the determinant of the square matrix obtained by eliminating the $i$th row and the $j$th column of A

$$M_{ij} = \det \mathrm{A}^{ij}$$

**Definition 46** (Cofactor)**.** The *cofactor* $\Delta_{ij}$ of the $ij$th element of a square matrix A is

$$\Delta_{ij} = (-1)^{i-j} M_{ij} = (-1)^{i-j} \det \mathrm{A}^{ij}$$

# 4 Matrix Inverses and Linear Equations

**Definition 47** (Inhomogeneous)**.** The system of equations $\mathrm{A}\mathbf{x} = \mathbf{d}$ is a system of *inhomogeneous* equations if $\mathbf{d} \neq \mathbf{0}$.

If $\mathbf{d} = \mathbf{0}$, it is a system of *homogeneous* equations.

**Definition 48** (Rank)**.** The *rank* of a matrix A, rank(A) is the maximum number of linearly independent columns/rows of A.

**Definition 49** (Singular)**.** A $n \times n$ square matrix A is *singular* if $\det \mathrm{A} = 0$ and *non-singular* if $\det \mathrm{A} \neq 0$.

# 5 Eigenvalues and Eigenvectors

**Definition 50** (Eigenvector & eigenvalue)**.** For a linear map $\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^n$, if $\mathcal{A}(\mathbf{x}) = \lambda \mathbf{x}$ for some *non-zero* vector $\mathbf{x} \in \mathbb{F}^n$ and $\lambda \in \mathbb{F}$, then $\mathbf{x}$ is an *eigenvector* of $\mathcal{A}$ with *eigenvalue* $\lambda$.

Similarly, for an $n \times n$ square matrix A, if $\mathrm{A}\mathbf{x} = \lambda \mathbf{x}$ for some *non-zero* vector $\mathbf{x} \in \mathbb{F}^n$ and $\lambda \in \mathbb{F}$, then $\mathbf{x}$ is an *eigenvector* of A with *eigenvalue* $\lambda$.

**Definition 51** (Characteristic polynomial)**.** The *characteristic polynomial* of the matrix A is the polynomial

$$p_{\mathrm{A}}(\lambda) = \det(\mathrm{A} - \lambda \mathrm{I})$$

The *characteristic equation* is $p_{\mathrm{A}}(\lambda) = 0$, and its solutions are the eigenvalues of A.

**Definition 52** (Algebraic multiplicity)**.** The multiplicity of an eigenvalue $\lambda$ as a root of the characteristic polynomial is called the *algebraic multiplicity* of $\lambda$, denoted $M_\lambda$. An eigenvalue with an algebraic multiplicity greater than one is *degenerate*.

**Definition 53** (Geometric multiplicity)**.** The maximum number, $m_\lambda$, of linearly independent eigenvectors corresponding to $\lambda$ is called the *geometric multiplicity* of $\lambda$.

**Definition 54** (Defect)**.** The difference $\Delta_\lambda = M_\lambda - m_\lambda$ is the *defect* of $\lambda$.

**Definition 55** (Generalised eigenvector)**.** A vector $\mathbf{x}$ is a *generalised eigenvector* of a map $\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^n$ if there is some eigenvalue $\lambda \in \mathbb{F}^n$ and some $k \in \mathbb{N}$ such that

$$(\mathcal{A} - \lambda \mathcal{I})^k(\mathbf{x}) = \mathbf{0}$$

**Definition 56** (Diagonal matrix)**.** A $n \times n$ matrix $\mathrm{D} = \{D_{ij}\}$ is a *diagonal matrix* if $D_{ij} = 0$ whenever $i \neq j$, i.e. if

$$\mathrm{D} = \begin{pmatrix} D_{11} & 0 & \dots & 0 \\ 0 & D_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D_{nn} \end{pmatrix}$$

**Definition 57** (Similar matrices)**.** The $n \times n$ matrices A and B are *similar*, or *conjugate*, if for some invertible matrix P,

$$\mathrm{B} = \mathrm{P}^{-1}\mathrm{AP}$$

**Definition 58** (Diagonalizable)**.** A linear map $\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^n$ is *diagonalizable* if $\mathbb{F}^n$ has a basis consisting of eigenvectors of $\mathcal{A}$.

An $n \times n$ matrix A is *diagonalizable* if A is similar with a diagonal matrix.

**Definition 59** (Normal)**.** A square matrix A is *normal* if

$$\mathrm{A}^\dagger\mathrm{A} = \mathrm{A}\mathrm{A}^\dagger$$

**Definition 60** (Form)**.** A map $\mathcal{F}(\mathbf{x})$

$$\mathcal{F}(\mathbf{x}) = \mathbf{x}^\dagger\mathrm{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^* A_{ij} x_j$$

is a *(sesquilinear) form*; A is its *coefficient matrix.*

**Definition 61** (Hermitian form)**.** If $\mathrm{A} = \mathrm{H}$ is an Hermitian matrix, the map $\mathcal{F}(\mathbf{x}) : \mathbb{C}^n \to \mathbb{C}$, where

$$\mathcal{F}(\mathbf{x}) = \mathbf{x}^\dagger\mathrm{H}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^* H_{ij} x_j$$

is an *Hermitian form* on $\mathbb{C}^n$.

**Definition 62** (Quadratic form)**.** If $\mathrm{A} = \mathrm{S}$ is a real symmetric matrix, the map $\mathcal{F}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$, where

$$\mathcal{F}(\mathbf{x}) = \mathbf{x}^\mathrm{T}\mathrm{S}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i S_{ij} x_j$$

is a *quadratic form* on $\mathbb{R}^n$.

**Definition 63** (Hessian matrix)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function such that all the second-order partial derivative exist. The *Hessian matrix* $\mathrm{H}(\mathbf{a})$ of $f$ at $\mathbf{a} \in \mathbb{R}^n$ is defined to have components

$$H_{ij}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$$