# MATH96023/MATH97032/MATH97140 - Computational Linear Algebra

***Edition 2020.0***

**Colin J. Cotter**

**Jul 08, 2020**

# CONTENTS

# CHAPTER ONE

# PRELIMINARIES

In this preliminary section we revise a few key linear algebra concepts that will be used in the rest of the course, emphasising the column space of matrices. We will quote some standard results that should be found in an undergraduate linear algebra course.

## 1.1 Matrices, vectors and matrix-vector multiplication

We will consider the multiplication of a vector

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad x_i \in \mathbb{C},\ i = 1, 2, \ldots, n, \text{ i.e. } x \in \mathbb{C}^n,$$

by a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

i.e. $A \in \mathbb{C}^{m\times n}$. $A$ has $m$ rows and $n$ columns so that the product

$$b = Ax$$

produces $b \in \mathbb{C}^m$, defined by

$$b_i = \sum_{j=1}^{n} a_{ij}x_j,\ i = 1, 2, \ldots, m.$$

In this course it is important to consider the general case where $m \neq n$, which has many applications in data analysis, curve fitting etc. We will usually state generalities in this course for vectors over the field $\mathbb{C}$, noting where things specialise to $\mathbb{R}$.

We can quickly check that the map $x \to Ax$ given by matrix multiplication is a linear map from $\mathbb{C}^n \to \mathbb{C}^m$, since it is straightforward to check from the definition that

$$A(\alpha x + y) = \alpha A x + A y,$$

for all $x, y \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$. (Exercise: show this for yourself.)

It is very useful to interpret matrix-vector multiplication as a linear combination of the columns of $A$ with coefficients taken from the entries of $x$. If we write $A$ in terms of the columns,

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix},$$

where

$$a_i \in \mathbb{C}^m, \, i = 1, 2, \dots, n,$$

then

$$b = \sum_{j=1}^{n} x_j a_j,$$

i.e. a linear combination of the columns of $A$ as described above.

We can extend this idea to matrix-matrix multiplication. Taking $A \in \mathbb{C}^{l\times m}$, $C \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{l\times n}$, with $B = AC$, then the components of $B$ are given by

$$b_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj}, \quad 1 \leq i \leq l, \, 1 \leq j \leq n.$$

Writing $b_j \in \mathbb{C}^m$ as the jth column of $B$, for $1 \leq j \leq n$, and $c_j$ as the jth column of $C$, we see that

$$b_j = A c_j.$$

This means that the jth column of $B$ is the matrix-vector product of $A$ with the jth column of $C$. This kind of "column thinking" is very useful in understanding computational linear algebra algorithms.

An important example is the outer product of two vectors, $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$. Here it is useful to see these vectors as matrices with one column, i.e. $u \in \mathbb{C}^{m\times 1}$ and $v \in \mathbb{C}^{n\times 1}$. The outer product is $uv^T \in \mathbb{C}^{m\times n}$. The columns of $v^T$ are just single numbers (i.e. vectors of length 1), so viewing this as a matrix multiplication we see

$$uv^T = \begin{pmatrix} uv_1 & uv_2 & \dots & uv_n \end{pmatrix},$$

which means that all the columns of $uv^T$ are multiples of $u$. We will see in the next section that this matrix has rank 1.

## 1.2 Range, nullspace and rank

In this section we'll quickly rattle through some definitions and results.

**Definition 1 (Range)** *The range of $A$, range$(A)$, is the set of vectors that can be expressed as $Ax$ for some $x$.*

The next theorem follows as a result of the column space interpretation of matrix-vector multiplication.

**Theorem 2** *range$(A)$ is the vector space spanned by the columns of $A$.*

**Definition 3 (Nullspace)** *The nullspace null$(A)$ of $A$ is the set of vectors $x$ satisfying $Ax = 0$, i.e.*

$$\mathit{null}(A) = \{x \in \mathbb{C}^n : Ax = 0\}.$$

**Definition 4 (Rank)** *The rank rank$(A)$ of $A$ is the dimension of the column space of $A$.*

If

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix},$$

the column space of $A$ is span$(a_1, a_2, \dots, a_n)$.

**Definition 5** *An $m \times n$ matrix $A$ is full rank if it has maximum possible rank i.e. rank equal to $\min(m, n)$.*

If $m \geq n$ then $A$ must have $n$ linearly independent columns to be full rank. The next theorem is then a consequence of the column space interpretation of matrix-vector multiplication.

**Theorem 6** *An $m \times n$ matrix $A$ is full rank if and only if it maps no two distinct vectors to the same vector.*

**Definition 7** *A matrix $A$ is called nonsingular, or invertible, if it is a square matrix ($m = n$) of full rank.*

## 1.3 Invertibility and inverses

This means that an invertible matrix has columns that form a basis for $\mathbb{C}^m$. Given the canonical basis vectors defined by

$$e_j = \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix},$$

i.e. $e_j$ has all entries zero except for the jth entry which is 1, we can write

$$e_j = \sum_{k=1}^{m} z_{ik} a_k, \quad 1 \leq j \leq m.$$

In other words,

$$\begin{aligned} I = \begin{pmatrix} e_1 & e_2 & \dots & e_m \end{pmatrix} \\ = ZA. \end{aligned}$$

We call $Z$ a (left) inverse of $A$. (Exercises: show that $Z$ is the unique left inverse of $A$, and show that $Z$ is also the unique right inverse of $A$, satisfying $I = AZ$.) We write $Z = A^{-1}$.

The first four parts of the next theorem are a consequence of what we have so far, and we shall quote the rest (see a linear algebra course).

**Theorem 8** *Let $A \in \mathbb{C}^{m \times m}$. Then the following are equivalent.*

1. *$A$ has an inverse.*
2. *rank$(A) = m$.*
3. *range$(A) = \mathbb{C}^m$.*
4. *null$(A) = \{0\}$.*
5. *0 is not an eigenvalue of $A$.*
6. *0 is not a singular value of $A$.*
7. *The determinant $\det(A) \neq 0$.*

Finding the inverse of a matrix can be seen as a change of basis. Considering the equation $Ax = b$, we have $x = A^{-1}b$ for invertible $A$. We have seen already that $b$ can be written as

$$b = \sum_{j=1}^{m} x_j a_j.$$

Since the columns of $A$ span $\mathbb{C}^m$, the entries of $x$ thus provide the unique expansion of $b$ in the columns of $A$ which form a basis. Hence, whilst the entries of $b$ give basis coefficients for $b$ in the canonical basis $(e_1, e_2, \ldots, e_m)$, the entries of $x$ give basis coefficients for $b$ in the basis given by the columns of $A$.

## 1.4 Orthogonal vectors and orthogonal matrices

**Definition 9 (Adjoint)** *The adjoint (or Hermitian conjugate) of $A \in \mathbb{C}^{m\times n}$ is a matrix $A^* \in \mathbb{C}^{n\times m}$ (sometimes written $A^\dagger$ or $A'$), with*

$$a^*_{ij} = \bar{a_{j}}_i,$$

*where the bar denotes the complex conjugate of a complex number. If $A^* = A$ then we say that $A$ is Hermitian.*

*For real matrices, $A^* = A^T$. If $A = A^T$, then we say that the matrix is symmetric.*

The following identity is very important when dealing with adjoints.

**Theorem 10** *For matrices $A$, $B$ with compatible dimensions (so that they can be multiplied),*

$$(AB)^* = B^*A^*.$$

## 1.5 Inner products and orthogonality

The inner product is a critical tool in computational linear algebra.

**Definition 11 (Inner product)** *Let $x, y \in \mathbb{C}^m$. Then the inner product of $x$ and $y$ is*

$$x^*y = \sum_{i=1}^{m} \bar{x}_i y_i.$$

(Exercise: check that the inner product is bilinear, i.e. linear in both of the arguments.)

We will frequently use the natural norm derived from the inner product to define size of vectors.

**Definition 12 (2-Norm)** *Let $x \in \mathbb{C}^m$. Then the 2-norm of $x$ is*

$$\|x\| = \sqrt{\sum_{i=1}^{m} x_i^2} = \sqrt{x^*x}.$$

Orthogonality will emerge as an early key concept in this course.

**Definition 13 (Orthogonal vectors)** *Let $x, y \in \mathbb{C}^m$. The two vectors are orthogonal if $x^* y = 0$.*

*Similarly, let $X$, $Y$ be two sets of vectors. The two sets are orthogonal if*

$$x^* y = 0 \, \forall x \in X, \; y \in Y.$$

*A set $S$ of vectors is itself orthogonal if*

$$x^* y = 0 \, \forall x, y \in S.$$

*We say that $S$ is orthonormal if we also have $\|x\| = 1$ for all $x \in S$.*

## 1.6 Orthogonal components of a vector

Let $S = \{q_1, q_2, \ldots, q_n\}$ be an orthonormal set of vectors in $\mathbb{C}^m$, and take another arbitrary vector $v \in \mathbb{C}^m$. Now take

$$r = v - (q_1^* v) q_1 - (q_2^* v) q_2 - \ldots (q_n^* v) q_n.$$

Then, we can check that $r$ is orthogonal to $S$, by calculating for each $1 \leq i \leq n$,

$$\begin{aligned} q_i^* r = q_i^* v - (q_1^* v)(q_i^* q_1) - \ldots (q_n^* v)(q_i^* q_n) \\ = q_i^* v - q_i^* v = 0, \end{aligned}$$

since $q_i^* q_j = 0$ if $i \neq j$, and 1 if $i = j$. Thus,

$$v = r + \sum_{i=1}^{n} (q_i^* v) q_i = r + \sum_{i=1}^{n} \underbrace{(q_i q_i^*)}_{\text{rank-1 matrix}} v.$$

If $S$ is a basis for $\mathbb{C}^m$, then $n = m$ and $r = 0$, and we have

$$v = \sum_{i=1}^{m} (q_i q_i^*) v.$$

## 1.7 Unitary matrices

**Definition 14** *A matrix $Q \in \mathbb{C}^{m \times m}$ is unitary if $Q^* = Q^{-1}$.*

*For real matrices, a matrix $Q$ is orthogonal if $Q^T = Q^{-1}$.*

**Theorem 15** *The columns of a unitary matrix $Q$ are orthonormal.*

**Proof 16** *We have $I = Q^* Q$. Then using the column space interpretation of matrix-matrix multiplication,*

$$e_j = Q^* q_j,$$

*where $q_j$ is the jth column of $Q$. Taking row i of $e_j$, we have*

$$\delta_{ij} = q_i^* q_j, \text{ where } \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{otherwise} \end{cases}.$$

Extending a theme from earlier, we can interpret $Q^* = Q^{-1}$ as representing a change of orthogonal basis. If $Qx = b$, then $x = Q^* b$ contains the coefficients of $b$ expanded in the basis given by the orthonormal columns of $Q$.

## 1.8 Vector norms

Various vector norms are useful to measure the size of a vector. In computational linear algebra we need them for quantifying errors etc.

**Definition 17** *A norm is a function $\|\cdot\|:\mathbb{C}^m\to\mathbb{R}$, such that*

1. $\|x\|\geq 0$, *and* $\|x\|=0 \implies x=0$.
2. $\|x+y\|\leq\|x\|+\|y\|$ *(triangle inequality).*
3. $\|\alpha x\|=|\alpha|\|x\|$ *for all* $x\in\mathbb{C}^m$ *and* $\alpha\in\mathbb{C}$.

We have already seen the 2-norm, or Euclidean norm, which is part of a larger class of norms called p-norms, with

$$\|x\|_p=\left(\sum_{i=1}^m |x_i|^p\right)^{1/p},$$

for real 'p>0'. We will also consider weighted norms

$$\|x\|_{W,p}=\|Wx\|_p,$$

where $W$ is a matrix.

## 1.9 Projectors and projections