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.

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

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

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,

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

where

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

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

$$b_j = Ac_j.$$

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

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

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 result follows directly from the column interpretation of matrix-vector multiplication.

Theorem 2 *The range of A is the vector space spanned by the columns of A.*

Definition 3 (Nullspace) *The nullspace of A, null(A), is the set of vectors $x \in \mathbb{C}^n$ such that $Ax = 0$.*

Definition 4 (Rank) *The rank of A, rank(A), is the dimension of the column space of A, i.e. span($a_1, a_2, \ldots, a_n$) where*

$$A = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix}.$$

Hence, we now understand why the outer product uv^T is rank 1.

Definition 5 (Full rank) *A matrix $A \in \mathbb{C}^{m\times n}$ is full rank if it has maximum possible rank (the minimum of m and n). If $m \geq n$, then A must have linearly independent columns to be full rank.*

We quote the following result which follows from the linear independence of of the columns of a full rank matrix A with $m \geq n$, and the column interpretation of matrix multiplication.

Theorem 6 *A matrix $A \in \mathbb{C}^{m\times n}$ with $m \geq n$ has full rank if and only if it maps no two distinct vectors to the same vector.*

Definition 7 (Invertible) *A matrix $A \in \mathbb{C}^{m\times n}$ is nonsingular (or equivalently, invertible) if it is square ($m = n$) and of full rank.*

A consequence of this definition is that the columns of A form a basis for $\mathbb{C}^m$.

Writing e_j as the jth canonical basis vector,

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

i.e. the vector with zeros everywhere except for in the jth entry, we can write

$$e_j = \sum_{i=1}^{m} z_{ij} a_i,$$

using the columns of A, since the columns form a basis, for coefficients z_{ij}, $1 \leq i, j \leq m$. In other words,

$$I = \begin{pmatrix} e_1 & e_2 & \ldots & e_m \end{pmatrix} = ZA,$$

and we call Z the (left) inverse of A.

Let A also have a right inverse Y so that $I = AY$. Then,

$$Z = ZI = ZAY = (ZA)Y = IY = Y,$$

so the left and right inverses agree, and we call them A^{-1}. (Exercise, show that a full-rank square matrix has unique left and right inverse.)

We quote the following result, which partially follows from what we have seen so far.