# Lecture 1. Matrix-Vector Multiplication

You already know the formula for matrix-vector multiplication. Nevertheless, the purpose of this first lecture is to describe a way of interpreting such products that may be less familiar. If $b = Ax$, then *$b$ is a linear combination of the columns of $A$.*

## Familiar Definitions

Let $x$ be an $n$-dimensional column vector and let $A$ be an $m \times n$ matrix ($m$ rows, $n$ columns). Then the matrix-vector product $b = Ax$ is an $m$-dimensional column vector defined as follows:

$$b_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, m. \tag{1.1}$$

Here $b_i$ denotes the $i$th entry of $b$, $a_{ij}$ denotes the $i,j$ entry of $A$ ($i$th row, $j$th column), and $x_j$ denotes the $j$th entry of $x$. For simplicity, we assume in all but a few lectures of this book that quantities such as these belong to $\mathbb{C}$, the field of complex numbers. The space of $m$-vectors is $\mathbb{C}^m$, and the space of $m \times n$ matrices is $\mathbb{C}^{m \times n}$.

The map $x \mapsto Ax$ is *linear*, which means that, for any $x, y \in \mathbb{C}^n$ and any $\alpha \in \mathbb{C}$,

$$\begin{aligned} A(x+y) &= Ax + Ay, \\ A(\alpha x) &= \alpha A x. \end{aligned}$$

Conversely, every linear map from $\mathbb{C}^n$ to $\mathbb{C}^m$ can be expressed as multiplication by an $m \times n$ matrix.

## A Matrix Times a Vector

Let $a_j$ denote the $j$th column of $A$, an $m$-vector. Then (1.1) can be rewritten

$$b = Ax = \sum_{j=1}^{n} x_j a_j. \tag{1.2}$$

This equation can be displayed schematically as follows:

$$\left[\begin{array}{c} \\ \\ b \\ \\ \\ \end{array}\right] = \left[\begin{array}{c|c|c|c} & & & \\ & & & \\ a_1 & a_2 & \cdots & a_n \\ & & & \\ & & & \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right] = x_1 \left[\begin{array}{c} \\ \\ a_1 \\ \\ \\ \end{array}\right] + x_2 \left[\begin{array}{c} \\ \\ a_2 \\ \\ \\ \end{array}\right] + \cdots + x_n \left[\begin{array}{c} \\ \\ a_n \\ \\ \\ \end{array}\right].$$

In (1.2), $b$ is expressed as a linear combination of the columns $a_j$. Nothing but a slight change of notation has occurred in going from (1.1) to (1.2). Yet thinking of $Ax$ in terms of the form (1.2) is essential for a proper understanding of the algorithms of numerical linear algebra.

One way to summarize these different ways of viewing matrix-vector products is like this. As mathematicians, we are used to viewing the formula $Ax = b$ as a statement that $A$ acts on $x$ to produce $b$. The formula (1.2), by contrast, suggests the interpretation that $x$ acts on $A$ to produce $b$.

**Example 1.1.** Fix a sequence of numbers $\{x_1, \ldots, x_m\}$. If $p$ and $q$ are polynomials of degree $< n$ and $\alpha$ is a scalar, then $p+q$ and $\alpha p$ are also polynomials of degree $< n$. Moreover, the values of these polynomials at the points $x_i$ satisfy the following linearity properties:

$$\begin{aligned} (p+q)(x_i) &= p(x_i) + q(x_i), \\ (\alpha p)(x_i) &= \alpha(p(x_i)). \end{aligned}$$

Thus the map from vectors of coefficients of polynomials $p$ of degree $< n$ to vectors $(p(x_1), p(x_2), \ldots, p(x_m))$ of sampled polynomial values is linear. Any linear map can be expressed as multiplication by a matrix; this is an example. In fact, it is expressed by an $m \times n$ *Vandermonde matrix*

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{bmatrix}.$$

If $c$ is the column vector of coefficients of $p$,

$$c = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{bmatrix}, \qquad p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1}x^{n-1},$$

then the product $Ac$ gives the sampled polynomial values. That is, for each $i$ from 1 to $m$, we have

$$(Ac)_i \;=\; c_0 + c_1 x_i + c_2 x_i^2 + \cdots + c_{n-1} x_i^{n-1} \;=\; p(x_i). \tag{1.3}$$

In this example, it is clear that the matrix-vector product $Ac$ need not be thought of as $m$ distinct scalar summations, each giving a different linear combination of the entries of $c$, as (1.1) might suggest. Instead, $A$ can be viewed as a matrix of columns, each giving sampled values of a monomial,

$$A = \left[\begin{array}{c|c|c|c|c} \\ 1 & x & x^2 & \cdots & x^{n-1} \\ \\ \end{array}\right], \tag{1.4}$$

and the product $Ac$ should be understood as a single vector summation in the form of (1.2) that at once gives a linear combination of these monomials,

$$Ac \;=\; c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1}x^{n-1} \;=\; p(x).$$

□

The remainder of this lecture will review some fundamental concepts in linear algebra from the point of view of (1.2).

## A Matrix Times a Matrix

For the matrix-matrix product $B = AC$, *each column of $B$ is a linear combination of the columns of $A$.* To derive this fact, we begin with the usual formula for matrix products. If $A$ is $\ell \times m$ and $C$ is $m \times n$, then $B$ is $\ell \times n$, with entries defined by

$$b_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj}. \tag{1.5}$$

Here $b_{ij}$, $a_{ik}$, and $c_{kj}$ are entries of $B$, $A$, and $C$, respectively. Written in terms of columns, the product is

$$\left[\begin{array}{c|c|c|c} \\ b_1 & b_2 & \cdots & b_n \\ \\ \end{array}\right] = \left[\begin{array}{c|c|c|c} \\ a_1 & a_2 & \cdots & a_m \\ \\ \end{array}\right] \left[\begin{array}{c|c|c|c} c_1 & c_2 & \cdots & c_n \\ \end{array}\right],$$

and (1.5) becomes

$$b_j = Ac_j = \sum_{k=1}^{m} c_{kj} a_k. \tag{1.6}$$

Thus $b_j$ is a linear combination of the columns $a_k$ with coefficients $c_{kj}$.

**Example 1.2.** A simple example of a matrix-matrix product is the *outer product*. This is the product of an $m$-dimensional column vector $u$ with an $n$-dimensional row vector $v$; the result is an $m \times n$ matrix of rank 1. The outer product can be written

$$\left[\begin{array}{c} \\ u \\ \\ \end{array}\right] \left[\begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array}\right] = \left[\begin{array}{c|c|c|c} \\ v_1 u & v_2 u & \cdots & v_n u \\ \\ \end{array}\right] = \left[\begin{array}{ccc} v_1 u_1 & \cdots & v_n u_1 \\ \vdots & & \vdots \\ v_1 u_m & \cdots & v_n u_m \end{array}\right].$$

The columns are all multiples of the same vector $u$, and similarly, the rows are all multiples of the same vector $v$. □

**Example 1.3.** As a second illustration, consider $B = AR$, where $R$ is the upper-triangular $n \times n$ matrix with entries $r_{ij} = 1$ for $i \le j$ and $r_{ij} = 0$ for $i > j$. This product can be written

$$\left[\begin{array}{c|c|c} \\ b_1 & \cdots & b_n \\ \\ \end{array}\right] = \left[\begin{array}{c|c|c} \\ a_1 & \cdots & a_n \\ \\ \end{array}\right] \left[\begin{array}{ccc} 1 & \cdots & 1 \\ & \ddots & \vdots \\ & & 1 \end{array}\right].$$

The column formula (1.6) now gives

$$b_j \;=\; Ar_j \;=\; \sum_{k=1}^{j} a_k. \tag{1.7}$$

That is, the $j$th column of $B$ is the sum of the first $j$ columns of $A$. The matrix $R$ is a discrete analogue of an indefinite integral operator. □

## Range and Nullspace

The *range* of a matrix $A$, written range$(A)$, is the set of vectors that can be expressed as $Ax$ for some $x$. The formula (1.2) leads naturally to the following characterization of range$(A)$:

**Theorem 1.1.** range$(A)$ *is the space spanned by the columns of* $A$.

*Proof.* By (1.2), any $Ax$ is a linear combination of the columns of $A$. Conversely, any vector $y$ in the space spanned by the columns of $A$ can be written as a linear combination of the columns, $y = \sum_{j=1}^{n} x_j a_j$. Forming a vector $x$ out of the coefficients $x_j$, we have $y = Ax$, and thus $y$ is in the range of $A$. □

In view of Theorem 1.1, the range of a matrix $A$ is also called the *column space* of $A$.

The *nullspace* of $A \in \mathbb{C}^{m\times n}$, written null($A$), is the set of vectors $x$ that satisfy $Ax = 0$, where 0 is the 0-vector in $\mathbb{C}^m$. The entries of each vector $x \in \text{null}(A)$ give the coefficients of an expansion of zero as a linear combination of columns of $A$: $0 = x_1a_1 + x_2a_2 + \cdots + x_na_n$.

## Rank

The *column rank* of a matrix is the dimension of its column space. Similarly, the *row rank* of a matrix is the dimension of the space spanned by its rows. Row rank always equals column rank (among other proofs, this is a corollary of the singular value decomposition, discussed in Lectures 4 and 5), so we refer to this number simply as the *rank* of a matrix.

An $m \times n$ matrix of *full rank* is one that has the maximal possible rank (the lesser of $m$ and $n$). This means that a matrix of full rank with $m \geq n$ must have $n$ linearly independent columns. Such a matrix can also be characterized by the property that the map it defines is one-to-one:

**Theorem 1.2.** *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.*

*Proof.* ($\Longrightarrow$) If $A$ is of full rank, its columns are linearly independent, so they form a basis for range($A$). This means that every $b \in \text{range}(A)$ has a unique linear expansion in terms of the columns of $A$, and therefore, by (1.2), every $b \in \text{range}(A)$ has a unique $x$ such that $b = Ax$. ($\Longleftarrow$) Conversely, if $A$ is not of full rank, its columns $a_j$ are dependent, and there is a nontrivial linear combination such that $\sum_{j=1}^{n} c_j a_j = 0$. The nonzero vector $c$ formed from the coefficients $c_j$ satisfies $Ac = 0$. But then $A$ maps distinct vectors to the same vector since, for any $x$, $Ax = A(x + c)$. □

## Inverse

A *nonsingular* or *invertible* matrix is a square matrix of full rank. Note that the $m$ columns of a nonsingular $m \times m$ matrix $A$ form a basis for the whole space $\mathbb{C}^m$. Therefore, we can uniquely express any vector as a linear combination of them. In particular, the canonical unit vector with 1 in the $j$th entry and zeros elsewhere, written $e_j$, can be expanded:

$$e_j = \sum_{i=1}^{m} z_{ij} a_i. \tag{1.8}$$

Let $Z$ be the matrix with entries $z_{ij}$, and let $z_j$ denote the $j$th column of $Z$. Then (1.8) can be written $e_j = Az_j$. This equation has the form (1.6); it can be written again, most concisely, as

$$\left[\begin{array}{c|c|c} \\ \\ e_1 & \cdots & e_m \\ \\ \\ \end{array}\right] = I = AZ.$$

The matrix $Z$ is the *inverse* of $A$. Any square nonsingular matrix $A$ has a unique inverse, written $A^{-1}$, that satisfies $AA^{-1} = A^{-1}A = I$.

The following theorem records a number of equivalent statements that hold when a square matrix is nonsingular. These conditions appear in linear algebra texts, and we shall not give a proof here. Concerning *(f)*, see Lecture 5.

**Theorem 1.3.** *For $A \in \mathbb{C}^{m\times m}$, the following conditions are equivalent:*

*(a) $A$ has an inverse $A^{-1}$,*
*(b)* $\operatorname{rank}(A) = m$,
*(c)* $\operatorname{range}(A) = \mathbb{C}^m$,
*(d)* $\operatorname{null}(A) = \{0\}$,
*(e) 0 is not an eigenvalue of $A$,*
*(f) 0 is not a singular value of $A$,*
*(g)* $\det(A) \neq 0$.

Concerning *(g)*, we mention that the determinant, though a convenient notion theoretically, rarely finds a useful role in numerical algorithms.

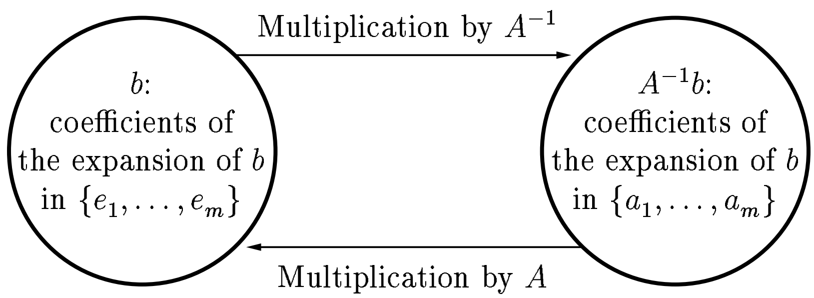
## A Matrix Inverse Times a Vector

When writing the product $x = A^{-1}b$, it is important not to let the inverse-matrix notation obscure what is really going on! Rather than thinking of $x$ as the result of applying $A^{-1}$ to $b$, we should understand it as the unique vector that satisfies the equation $Ax = b$. By (1.2), this means that $x$ is the vector of coefficients of the unique linear expansion of $b$ in the basis of columns of $A$.

This point cannot be emphasized too much, so we repeat:

*$A^{-1}b$ is the vector of coefficients of the expansion of $b$*
*in the basis of columns of $A$.*

Multiplication by $A^{-1}$ is a *change of basis* operation:




In this description we are being casual with terminology, using "$b$" in one instance to denote an $m$-tuple of numbers and in another as a point in an abstract vector space. The reader should think about these matters until he or she is comfortable with the distinction.

## A Note on $m$ and $n$

Throughout numerical linear algebra, it is customary to take a rectangular matrix to have dimensions $m \times n$. We follow this convention in this book.

What if the matrix is square? The usual convention is to give it dimensions $n \times n$, but in this book we shall generally take the other choice, $m \times m$. Many of our algorithms require us to look at rectangular submatrices formed by taking a subset of the columns of a square matrix. If the submatrix is to be $m \times n$, the original matrix had better be $m \times m$.

## Exercises

1. Let $B$ be a $4 \times 4$ matrix to which we apply the following operations:
    1. double column 1,
    2. halve row 3,
    3. add row 3 to row 1,
    4. interchange columns 1 and 4,
    5. subtract row 2 from each of the other rows,
    6. replace column 4 by column 3,
    7. delete column 1 (so that the column dimension is reduced by 1).

    (a) Write the result as a product of eight matrices.

    (b) Write it again as a product $ABC$ (same $B$) of three matrices.

2. Suppose masses $m_1, m_2, m_3, m_4$ are located at positions $x_1, x_2, x_3, x_4$ in a line and connected by springs with spring constants $k_{12}, k_{23}, k_{34}$ whose natural

lengths of extension are $\ell_{12}, \ell_{23}, \ell_{34}$. Let $f_1, f_2, f_3, f_4$ denote the rightward forces on the masses, e.g., $f_1 = k_{12}(x_2 - x_1 - \ell_{12})$.

(a) Write the $4 \times 4$ matrix equation relating the column vectors $f$ and $x$. Let $K$ denote the matrix in this equation.

(b) What are the dimensions of the entries of $K$ in the physics sense (e.g., mass times time, distance divided by mass, etc.)?

(c) What are the dimensions of $\det(K)$, again in the physics sense?

(d) Suppose $K$ is given numerical values based on the units meters, kilograms, and seconds. Now the system is rewritten with a matrix $K'$ based on centimeters, grams, and seconds. What is the relationship of $K'$ to $K$? What is the relationship of $\det(K')$ to $\det(K)$?

3. Generalizing Example 1.3, we say that a square or rectangular matrix $R$ with entries $r_{ij}$ is *upper-triangular* if $r_{ij} = 0$ for $i > j$. By considering what space is spanned by the first $n$ columns of $R$ and using (1.8), show that if $R$ is a nonsingular $m \times m$ upper-triangular matrix, then $R^{-1}$ is also upper-triangular. (The analogous result also holds for lower-triangular matrices.)

4. Let $f_1, \ldots, f_8$ be a set of functions defined on the interval $[1, 8]$ with the property that for any numbers $d_1, \ldots, d_8$, there exists a set of coefficients $c_1, \ldots, c_8$ such that

$$\sum_{j=1}^{8} c_j f_j(i) = d_i, \qquad i = 1, \ldots, 8.$$

(a) Show by appealing to the theorems of this lecture that $d_1, \ldots, d_8$ determine $c_1, \ldots, c_8$ uniquely.

(b) Let $A$ be the $8 \times 8$ matrix representing the linear mapping from data $d_1, \ldots, d_8$ to coefficients $c_1, \ldots, c_8$. What is the $i, j$ entry of $A^{-1}$?