# Matrix Operations

In this appendix we list some of the important facts about matrix operations and solutions to systems of linear equations.

## A.1. Matrix Multiplication

The product of a row $\mathbf{a} = (a_1, \ldots, a_n)$ and a column $\mathbf{x} = (x_1, \ldots, x_n)^T$ is a scalar:

$$\mathbf{a}\,\mathbf{x} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1 + \cdots + a_nx_n = x_1a_1 + \cdots + x_na_n. \quad \text{(A.1)}$$

The product of an $m \times n$ *matrix* $A$ and the column vector $\mathbf{x}$ has two definitions, and you should check that they are equivalent. If we think of $A$ as being made of $m$ rows $\mathbf{r}_i$, then

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{r}_1\mathbf{x} \\ \mathbf{r}_2\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix}. \quad \text{(A.2)}$$

In practice, that is how the product $A\mathbf{x}$ is usually calculated. However, it is often better to think of $A$ as being comprised of $n$ columns $\mathbf{a}_i$, each of height $m$. From that perspective,

$$A\mathbf{x} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n. \qquad \text{(A.3)}$$

That is, the product of a matrix with a vector is a linear combination of the columns of the vector, with the entries of the vector providing the coefficients.

Finally, we consider the product of two matrices. If $A$ is an $m \times n$ matrix and $B$ is an $n \times p$ matrix, then $AB$ is an $m \times p$ matrix whose $ij$ entry is the product of the $i^{\text{th}}$ row of $A$ and the $j^{\text{th}}$ column of $B$. That is,

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}. \qquad \text{(A.4)}$$

This can also be expressed in terms of the columns of $B$.

$$AB = A(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p) = (A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_p). \qquad \text{(A.5)}$$

The matrix $A$ acts separately on each column of $B$.

## A.2. Row reduction

The three standard row operations are:

(1) Multiplying a row by a nonzero scalar.
(2) Adding a multiple of one row to another.
(3) Swapping the positions of two rows.

Each of these steps is reversible, so if you can get from $A$ to $B$ by row operations, then you can also get from $B$ to $A$. In that case we say that the matrices $A$ and $B$ are *row-equivalent*.

**Definition.** *A matrix is said to be in* row-echelon form *if (1) any rows made completely of zeroes lie at the bottom of the matrix and (2) the first nonzero entries of the various rows form a staircase pattern: the first nonzero entry of the $k+1^{st}$ row is to the right of the first nonzero entry of the $k^{th}$ row.*

For instance, of the matrices

$$\begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}, \qquad \text{(A.6)}$$

only the first is in row-echelon form. In the second matrix, a row of zeroes lies above a nonzero row. In the third matrix, the first nonzero entry of the third row is under, not to the right of, the first nonzero entry of the second row.

**Definition.** *If a matrix is in row-echelon form, then the first nonzero entry of each row is called a* pivot*, and the columns in which pivots appear are called* pivot columns*.*

If two matrices in row-echelon form are row-equivalent, then their pivots are in exactly the same places. When we speak of the pivot columns of a general matrix $A$, we mean the pivot columns of any matrix in row-echelon form that is row-equivalent to $A$.

It is always possible to convert a matrix to row-echelon form. The standard algorithm is called *Gaussian elimination* or *row reduction*. Here it is applied to the matrix

$$A = \begin{pmatrix} 2 & -2 & 4 & -2 \\ 2 & 1 & 10 & 7 \\ -4 & 4 & -8 & 4 \\ 4 & -1 & 14 & 6 \end{pmatrix}. \tag{A.7}$$

(1) Subtract the first row from the second.
(2) Add twice the first row to the third.
(3) Substract twice the first row from the fourth. At this point the matrix is

$$\begin{pmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 6 & 10 \end{pmatrix}. \tag{A.8}$$

(4) Subtract the second row from the fourth.
(5) Finally, swap the third and fourth rows. This gives a matrix,

$$A_{\text{ref}} = \begin{pmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{A.9}$$

in row-echelon form, that is row-equivalent to $A$. To get a particularly nice form, we can continue to do row operations:

(6) Divide the first row by 2.
(7) Divide the second row by 3.
(8) Add the third row to the first.
(9) Subtract three times the third row from the second.
(10) Add the second row to the first.

This gives a matrix,

$$A_{\text{rref}} = \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{A.10}$$

in what is called *reduced row-echelon form.*

**Definition.** *A matrix is in* reduced row-echelon form *if (1) it is in row-echelon form, (2) all of the pivots are equal to 1, and (3) all entries in the pivot columns, except for the pivots themselves, are equal to zero.*

For any matrix $A$ there is a unique matrix $A_{\text{rref}}$, in reduced row-echelon form, that is row-equivalent to $A$. $A_{\text{rref}}$ is called *the reduced row-echelon form of $A$.* Most computer linear algebra programs have a built-in routine for converting a matrix to reduced row-echelon form. In MATLAB it is "rref".

## A.3. Rank

**Definition.** *The* rank *of a matrix is the number of pivots in its reduced row-echelon form.*

Note that the rank of an $m \times n$ matrix cannot be bigger than $m$, since you can't have more than one pivot per row. It also can't be bigger than $n$, since you can't have more than one pivot per column. If $m < n$, then the rank is always less than $n$ and there are at least $n - m$ columns without pivots. If $m > n$, then the rank is always less than $m$ and there are at least $m - n$ rows of zeroes in the reduced row-echelon form.

If we have a square $n \times n$ matrix, then either the rank equals $n$, in which case the reduced row-echelon form is the identity matrix, or the rank is less than $n$, in which case there is a row of zeroes in the reduced row-echelon form, *and* there is at least one column without a pivot. In the first case we say the matrix is *invertible*, and in the second case we say the matrix is *singular*. The determinant of the matrix tells the difference between the two cases. The determinant of a singular matrix is always zero, while the determinant of an invertible matrix is always nonzero.

As we shall soon see, the rank of a matrix equals the dimension of its column space. A basis for the column space can be deduced from the positions of the pivots. The dimension of the null space of a matrix equals the number of columns without pivots, namely $n$ minus the rank, and a basis for the null space can be deduced from the reduced row-echelon form of the matrix.

## A.4. Solving $A\mathbf{x} = 0$.

Suppose we are given a system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned} \tag{A.11}$$

This is more easily written as $A\mathbf{x} = 0$, where $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$. Since multiplying equations by nonzero constants, adding equations and swapping the order of equations doesn't change the solution, the equations $A\mathbf{x} = 0$ are equivalent to $A_{\text{ref}}\mathbf{x} = 0$, or to $A_{\text{rref}}\mathbf{x} = 0$. Solving $A\mathbf{x} = 0$ boils down to two steps:

(1) Putting $A$ in reduced row-echelon form. This can be done by Gaussian elimination or by computer.

(2) Understanding how to read off the solutions to $A_{\text{rref}}\mathbf{x} = 0$ from the entries of $A_{\text{rref}}$.

As an example, consider the matrix $A_{\text{rref}}$ of equation (A.10). The four equations read:

$$\begin{aligned} x_1 + 4x_3 &= 0, \\ x_2 + 2x_3 &= 0, \\ x_4 &= 0, \\ 0 &= 0. \end{aligned} \tag{A.12}$$

Since there are pivots in the first, second and fourth columns, we call $x_1$, $x_2$ and $x_4$ *pivot variables*, or *constrained variables*. The remaining variable, $x_3$, is called *free*. Each nontrivial equation involves exactly one of the constrained variables. They give that variable in terms of the free variables. Adding the trivial equation $x_3 = x_3$ and removing the $0 = 0$ equation we get:

$$\begin{aligned} x_1 &= -4x_3, \\ x_2 &= -2x_3, \\ x_3 &= x_3, \\ x_4 &= 0. \end{aligned} \tag{A.13}$$

In other words, the free variable $x_3$ can be whatever we wish, and determines our entire solution. The set of solutions to $A\mathbf{x} = 0$, also known as the null space of $A$ or the kernel of $A$, is all multiples of $(-4, -2, 1, 0)^T$.

As a second example, consider the matrix

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 8 & 11 & 14 \\ 1 & 3 & 5 & 8 & 11 \\ 4 & 10 & 16 & 23 & 30 \end{pmatrix}, \tag{A.14}$$

whose reduced row-echelon form is

$$B_{\text{rref}} = \begin{pmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{A.15}$$

The constrained (pivot) variables are $x_1$, $x_2$ and $x_4$, while $x_3$ and $x_5$ are free. The equations $B\mathbf{x} = 0$ are equivalent to $B_{\text{rref}}\mathbf{x} = 0$, which read:

$$\begin{aligned} x_1 &= x_3 - x_5, \\ x_2 &= -2x_3 + 2x_5, \\ x_4 &= \quad - 2x_5. \end{aligned} \tag{A.16}$$

Throwing in the dummy equations $x_3 = x_3$ and $x_5 = x_5$, we get

$$\mathbf{x} = x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \\ 1 \end{pmatrix}. \tag{A.17}$$

The null space of $B$ (i.e., the set of solutions to $B\mathbf{x} = 0$) is 2-dimensional, with a basis given by the solution with with $x_3 = 1$ and $x_5 = 0$ and the solution with $x_3 = 0$ and $x_5 = 1$, namely $\{(1, -2, 1, 0, 0)^T, (-1, 2, 0, -1, 1)^T\}$.

In general, the dimension of the null space is the number of free variables. The basis vectors are obtained by setting one of the free variables equal to one, setting the others equal to zero, and using the equations $B_{\text{rref}}\mathbf{x} = 0$ to solve for the constrained variables.

## A.5. The column space

The column space of a matrix is the span of its columns. This is equal to the span of the pivot columns. The pivot columns are themselves linearly independent, and so form a basis for the column space.

For example, if $B$ is as in (A.14), then the pivot columns are the first, second and fourth, as can be read off from the reduced row-echelon form (A.15). This means that the column space of $B$ is 3-dimensional, and that a basis is given by $\{(1, 2, 1, 4)^T, (2, 5, 3, 10)^T, (4, 11, 8, 23)^T\}$. Note that we do *not* use the columns of $B_{\text{rref}}$! We use the columns of $B$.

Let $\mathbf{b}_1, \ldots, \mathbf{b}_5$ be the columns of $B$. The equations $B\mathbf{x} = 0$ are the same as

$$x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3 + x_4\mathbf{b}_4 + x_5\mathbf{b}_5 = 0. \tag{A.18}$$

Since there is a solution with $x_3 = 1$ and $x_5 = 0$, we can write $\mathbf{b}_3$ as a linear combination of $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_4$, specifically $\mathbf{b}_3 = -\mathbf{b}_1 + 2\mathbf{b}_2$. Likewise, we can write $\mathbf{b}_5 = \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_4$. Since the columns corresponding to the free variables are linear combinations of the pivot columns, the span of the pivot columns is the entire column space. To see that the pivot columns are linearly independent, suppose that $x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_4\mathbf{b}_4 = 0$. This is a solution to $B\mathbf{x} = 0$ with $x_3 = x_5 = 0$. However, by (A.17), this means that $x_1 = x_2 = x_4 = 0$.

The same arguments apply to any matrix $A$. Each solution to $A\mathbf{x} = 0$ with one free variable equal to 1 and the rest equal to zero shows that the column corresponding to the free variable is a linear combination of the pivot columns. This means that the pivot columns span the column space of $A$. Since the only solution to $A\mathbf{x}$ with all the free variables zero is $\mathbf{x} = 0$, the pivot columns are linearly independent.

Finally, when do the columns of an $m \times n$ matrix span $\mathbb{R}^m$? We already know that the dimension of the column space equals the rank of the matrix. If this is $m$, then the columns span $\mathbb{R}^m$. If the rank is less than $m$, then the columns do not span $\mathbb{R}^m$.

## A.6. Summary

(1) The product of a matrix and a column vector can be viewed in two ways, either by multiplying the rows of the matrix by the vector, or by taking a linear combination of the columns of the matrix with coefficients given by the entries of the vector.

(2) Using row operations, we can convert any matrix $A$ into a reduced row-echelon form $A_{\mathrm{rref}}$. This form is unique.

(3) The rank of a matrix is the number of pivots in its reduced row-echelon form. This equals the dimension of the column space. The pivot columns of $A$ (not of $A_{\mathrm{rref}}$!) are a basis for the column space of $A$.

(4) The columns of an $m \times n$ matrix $A$ span $\mathbb{R}^m$ if and only if $A_{\mathrm{rref}}$ has a pivot in each row, or equivantly if the rank of $A$ equals $m$.

(5) The solutions to $A\mathbf{x} = 0$ are the same as the solutions to $A_{\mathrm{rref}}\mathbf{x} = 0$. Those equations give the pivot variables in terms of the free variables. The dimension of the null space of $A$ equals the number of free variables.

(6) The $m \times n$ matrix $A$ is 1–1 if and only if $A_{\text{rref}}$ has a pivot in each column, or equivalently if the rank equals $n$.

(7) A square $n \times n$ matrix either has rank $n$, in which case its determinant is nonzero, it is row-equivalent to the identity, its columns are linearly independent, and its columns span $\mathbb{R}^n$, or it has rank less than $n$, in which case its determinant is zero, its columns are linearly dependent, and its columns fail to span $\mathbb{R}^n$.