9. Properties of Matrices

Block Matrices

It is often convenient to partition a matrix M into smaller matrices called *blocks*, like so:

$$M = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 1 \\ \hline 0 & 1 & 2 & 0 \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

Here $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 1 & 2 \end{pmatrix}$, $D = (0)$.

- The blocks of a block matrix must fit together to form a rectangle. So $\left(\begin{array}{c|c} B & A \\ \hline D & C \end{array}\right)$ makes sense, but $\left(\begin{array}{c|c} C & B \\ \hline D & A \end{array}\right)$ does not.

- There are many ways to cut up an $n \times n$ matrix into blocks. Often context or the entries of the matrix will suggest a useful way to divide the matrix into blocks. For example, if there are large blocks of zeros in a matrix, or blocks that look like an identity matrix, it can be useful to partition the matrix accordingly.

- Matrix operations on block matrices can be carried out by treating the blocks as matrix entries. In the example above,

$$\begin{aligned} M^2 &= \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \\ &= \left(\begin{array}{c|c} A^2 + BC & AB + BD \\ \hline CA + DC & CB + D^2 \end{array}\right) \end{aligned}$$

Computing the individual blocks, we get:

$$\begin{aligned}
A^2 + BC &= \begin{pmatrix} 30 & 37 & 44 \\ 66 & 81 & 96 \\ 102 & 127 & 152 \end{pmatrix} \\
AB + BD &= \begin{pmatrix} 4 \\ 10 \\ 16 \end{pmatrix} \\
CA + DC &= \begin{pmatrix} 18 \\ 21 \\ 24 \end{pmatrix} \\
CB + D^2 &= (2)
\end{aligned}$$

Assembling these pieces into a block matrix gives:

$$\left(\begin{array}{ccc|c} 30 & 37 & 44 & 4 \\ 66 & 81 & 96 & 10 \\ 102 & 127 & 152 & 16 \\ \hline 4 & 10 & 16 & 2 \end{array}\right)$$

This is exactly M^2.

The Algebra of Square Matrices

Not every pair of matrices can be multiplied. When multiplying two matrices, the number of rows in the left matrix must equal the number of columns in the right. For an $r \times k$ matrix M and an $s \times l$ matrix N, then we must have $k = s$.

This is not a problem for square matrices of the same size, though. Two $n \times n$ matrices can be multiplied in either order. For a single matrix $M \in M_n^n$, we can form $M^2 = MM$, $M^3 = MMM$, and so on, and define $M^0 = I_n$, the identity matrix.

As a result, any polynomial equation can be evaluated on a matrix.

Example Let $f(x) = x - 2x^2 + 3x^3$.

Let $M = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$. Then:

$$M^2 = \begin{pmatrix} 1 & 2t \\ 0 & 1 \end{pmatrix}, M^3 = \begin{pmatrix} 1 & 3t \\ 0 & 1 \end{pmatrix}, \ldots$$

Then:

$$\begin{aligned} f(M) &= \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} - 2 \begin{pmatrix} 1 & 2t \\ 0 & 1 \end{pmatrix} + 3 \begin{pmatrix} 1 & 3t \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 6t \\ 0 & 2 \end{pmatrix} \end{aligned}$$

Suppose $f(x)$ is any function defined by a convergent Taylor Series:

$$f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \dots$$

Then we can define the matrix function by just plugging in M:

$$f(M) = f(0) + f'(0)M + \frac{1}{2!}f''(0)M^2 + \dots$$

There are additional techniques to determine the convergence of Taylor Series of matrices, based on the fact that the convergence problem is simple for diagonal matrices. It also turns out that $\exp(M) = 1+M+\frac{1}{2}M^2+\frac{1}{3!}M^3+\dots$ always converges.

Matrix multiplication does *not* commute. For generic $n \times n$ square matrices M and N, then $MN \neq NM$. For example:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

On the other hand:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Since $n \times n$ matrices are linear transformations $\mathbb{R}^n \to \mathbb{R}^n$, we can see that the order of successive linear transformations matters. For two linear transformations K and L taking $\mathbb{R}^n \to \mathbb{R}^n$, and $v \in \mathbb{R}^n$, then in general $K(L(v)) \neq L(K(v))$!

Finding matrices such that $MN = NM$ is an important problem in mathematics.

Trace

Matrices contain a great deal of information, so finding ways to extract essential information is useful.

Definition The *trace* of a square matrice $M = (m^i_j)$ is the sum of its diagonal entries.

$$\operatorname{tr} M = \sum_{i=1}^{n} m^i_i$$

Example

$$\operatorname{tr}\begin{pmatrix} 2 & 7 & 6 \\ 9 & 5 & 1 \\ 4 & 3 & 8 \end{pmatrix} = 2+5+8 = 15$$

While matrix multiplication does not commute, the trace of a product of matrices does not depend on the order of multiplication:

$$\begin{aligned}
\operatorname{tr}(MN) &= \operatorname{tr}(\sum_l M^i_l N^l_j) \\
&= \sum_i \sum_l M^i_l N^l_i \\
&= \sum_l \sum_i N^l_i M^i_l \\
&= \operatorname{tr}(\sum_i N^l_i M^i_l) \\
&= \operatorname{tr}(NM).
\end{aligned}$$

Thus, $\operatorname{tr}(MN) = \operatorname{tr}(NM)$ for any square matrices M and N.

In the previous example,

$$M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

$$MN = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq NM = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

However, $\operatorname{tr}(MN) = 2+1 = 3 = 1+2 = \operatorname{tr}(NM)$.

Another useful property of the trace is that:

$$\operatorname{tr} M = \operatorname{tr} M^T$$

This is true because the trace only uses the diagonal entries, which are fixed by the transpose. For example: $\operatorname{tr}\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = 4 = \operatorname{tr}\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \operatorname{tr}\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}^T$

Finally, trace is a linear transformation from matrices to the real numbers. This is easy to check.

Linear Systems Redux Recall that we can view a linear system as a matrix equation

$$MX = V,$$

with M an $r \times k$ matrix of coefficients, x a $k \times 1$ matrix of unknowns, and V an $r \times 1$ matrix of constants. If M is a square matrix, then the number of equations (r) is the same as the number of unknowns (k), so we have hope of finding a single solution.

Above we discussed functions of matrices. An extremely useful function would be $f(M) = \frac{1}{M}$, where $M\frac{1}{M} = I$. If we could compute $\frac{1}{M}$, then we would multiply both sides of the equation $MX = V$ by $\frac{1}{M}$ to obtain the solution immediately: $X = \frac{1}{M}V$.

Clearly, if the linear system has no solution, then there can be no hope of finding $\frac{1}{M}$, since if it existed we could find a solution. On the other hand, if the system has more than one solution, it also seems unlikely that $\frac{1}{M}$ would exist, since $X = \frac{1}{M}V$ yields only a single solution.

Therefore $\frac{1}{M}$ only sometimes exists. It is called the *inverse* of M, and is usually written M^{-1}.

References

Wikipedia:

- Trace (Linear Algebra)
- Block Matrix

Review Questions

1. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}$. Find AA^T and A^TA. What can you say about matrices MM^T and M^TM in general? Explain.

2. Compute $\exp(A)$ for the following matrices:

   - $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$
   - $A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$

- $A = \begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix}$

3. Suppose $ad - bc \neq 0$, and let $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

   (a) Find a matrix M^{-1} such that $MM^{-1} = I$.

   (b) Explain why your result explains what you found in a previous homework exercise.

   (c) Compute $M^{-1}M$.

4. Let $M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$. Divide M into named blocks, and then multiply blocks to compute M^2.