Lecture 12

Definition 12.1. A matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ is called *symmetric* if $A = A^T$, for example $\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$. A square matrix is called *diagonal* if all its off-diagonal entries are zero, for example:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Note that a diagonal matrix is always symmetric.

An application of matrices

Example 12.2. Suppose I want to compute the nutritional value of what someone eats throughout a week. We can encode the data in matrices, for easy computation:

$$\text{Nutrients: } D \;=\; \begin{array}{c} \begin{array}{cc} \scriptstyle\text{pizza} & \scriptstyle\text{beer} \end{array} \\ \begin{pmatrix} 282 & 140 \\ 13 & 0 \\ 7 & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \scriptstyle\text{kcal} \\ \scriptstyle\text{fat} \\ \scriptstyle\text{protein} \end{array},$$

$$\text{Quantity: } W = \begin{array}{c} \begin{array}{ccccccc} \scriptstyle M & \scriptstyle T & \scriptstyle W & \scriptstyle Th & \scriptstyle F & \scriptstyle Sa & \scriptstyle Su \end{array} \\ \begin{pmatrix} 2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 5 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} \\ \scriptstyle\text{pizza} \\ \scriptstyle\text{beer} \end{array}.$$

If we take the matrix product $P = DW$, we get a 3×7 matrix

$$P = (p_{ij}) = \begin{array}{c} \begin{array}{ccccccc} \scriptstyle M & \scriptstyle T & \scriptstyle W & \scriptstyle Th & \scriptstyle F & \scriptstyle Sa & \scriptstyle Su \end{array} \\ \begin{pmatrix} 704 & 0 & 0 & 0 & 982 & 0 & 0 \\ 26 & 0 & 0 & 0 & 13 & 0 & 0 \\ 15 & 0 & 0 & 0 & 12 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} \\ \scriptstyle\text{kcal} \\ \scriptstyle\text{fat} \\ \scriptstyle\text{protein} \end{array}.$$

So, for example, the total calorie intake on Monday is $p_{11} = 704$ and the total fat intake on Friday is $p_{25} = 13$.

The point is that the matrix D always stays as a constant, even if the weekly eating habits change, so this is a convenient way of encoding and computing such data. The matrix P encodes everything we want in a neat form.

Matrices and series

Let A be an $m \times m$ matrix (i.e., a square one). For any integer $n = 1, 2, 3, \ldots$ we can compute the nth power of A:

$$A^n = \underbrace{A \cdot A \cdots A}_{n \text{ times}}.$$

Moreover, we define $A^0 = I_m = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, which is called the *identity matrix* of size m.

So, if we have a polynomial $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$, then we can evaluate it on the matrix A:

$$p(A) = c_0I_m + c_1A + c_2A^2 + \cdots + c_nA^n.$$

The value $p(A)$ is still an $m \times m$ matrix.

Now, let $f(x) = \sum_{n=0}^{\infty} c_nx^n$ be a power series. We define

$$f(A) = \sum_{n=0}^{\infty} c_nA^n.$$

This series may or may not converge to a matrix.

Example 12.3. Let $A = \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}$. Then

$$A^2 = \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & xy \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$A^3 = \begin{pmatrix} 0 & 0 & xy \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0.$$

So any higher power of A is also 0. Thus, we can compute any power series of A, for example, the exponent:

$$\begin{aligned} \exp(A) = e^A = \sum_{n=0}^{\infty} \frac{1}{n!}A^n &= I_3 + A + \frac{1}{2}A^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & xy/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & x & xy/2 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Example 12.4. If we have a diagonal matrix, things are easy. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Then

$$A^2 = \begin{pmatrix} 1^2 & 0 \\ 0 & 2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 1 & 0 \\ 0 & 2^3 \end{pmatrix}, \ldots$$

so we have $e^A = \sum_{n=0}^{\infty} \frac{1}{n!}A^n = \begin{pmatrix} e^1 & 0 \\ 0 & e^2 \end{pmatrix}$.

Systems of linear equations

We now return to one of the original motivations for matrices. We will explain how matrices can be used to efficiently solve systems of linear equations.

Example 12.5. We solve system by successive transformations:

$$\left\{\begin{array}{rcr} y+3z &=& -1 \\ x+y+2z &=& 1 \\ 2x+y &=& 2 \end{array}\right. \xrightarrow{R_1 \leftrightarrow R_2} \left\{\begin{array}{rcr} x+y+2z &=& 1 \\ y+3z &=& -1 \\ 2x+y &=& 2 \end{array}\right. \quad \begin{array}{c}(R_i \text{ denotes row } i) \\ \text{(the first step just swaps rows 1 and 2)}\end{array}$$

$$\xrightarrow{R_3-2R_1} \left\{\begin{array}{rcr} x+y+2z &=& 1 \\ y+3z &=& -1 \\ -y-4z &=& 0 \end{array}\right. \text{ (eliminates } x \text{ from row 3)}$$

$$\xrightarrow{R_3+R_2} \left\{\begin{array}{rcr} x+y+2z &=& 1 \\ y+3z &=& -1 \\ -z &=& -1 \end{array}\right. \text{ (eliminates } y \text{ from row 3)} \xrightarrow{-R_3} \left\{\begin{array}{rcr} x+y+2z &=& 1 \\ y+3z &=& -1 \\ z &=& 1 \end{array}\right.$$

$$\xrightarrow{R_2-3R_3} \left\{\begin{array}{rcr} x+y+2z &=& 1 \\ y &=& -4 \\ z &=& 1 \end{array}\right. \text{ (eliminates } z \text{ from row 2)}$$

$$\xrightarrow{R_1-2R_3} \left\{\begin{array}{rcr} x+y &=& -1 \\ y &=& -4 \\ z &=& 1 \end{array}\right. \text{ (eliminates } z \text{ from row 1)}$$

$$\xrightarrow{R_1-R_2} \left\{\begin{array}{rcr} x &=& 3 \\ y &=& -4 \\ z &=& 1 \end{array}\right. \text{ (eliminates } y \text{ from row 1).}$$

To describe a streamlined algorithm for solving systems, we reformulate the problem in terms of matrices. The system above can be written

$$\begin{pmatrix} 0 & 1 & 3 \\ 1 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

If we let A be the 3×3 matrix, $\underline{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\underline{b} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$, then we can just write this as $A\underline{x} = \underline{b}$.

What did we do when we solved the system in (12.5)? We only transformed A and $\underline{b}$ in seven steps:

Step 1 Swap two eqns/rows.

Steps 2,3,5,6,7 Add a multiple of an eqn to another eqn.

Step 4 Multiply an eqn by a number.

When we do something to the equations, we can do the same to the rows of the matrices A and $\underline{b}$. So, we create the *augmented matrix* of the system

$$(A|\underline{b}) = \left(\begin{array}{ccc|c} 0 & 1 & 3 & -1 \\ 1 & 1 & 2 & 1 \\ 2 & 1 & 0 & 2 \end{array}\right).$$

We can now perform the seven steps above in terms of this matrix alone. For example, Steps 1-2 would be

$$\left(\begin{array}{ccc|c} 0 & 1 & 3 & -1 \\ 1 & 1 & 2 & 1 \\ 2 & 1 & 0 & 2 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 3 & -1 \\ 2 & 1 & 0 & 2 \end{array}\right) \xrightarrow{R_3 - 2R_1} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 3 & -1 \\ 0 & -1 & -4 & 0 \end{array}\right).$$

Elementary row operations

The operations we have used in Steps 1-7 above are called *elementary row operations (ERO).* The imporant thing to note is that they do *not* change the solutions to the system. The goal is to obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

since this gives us $x = 3$, $y = -4$, $z = 1$.

We will now use matrices to solve another system:

Example 12.6. Solve the system

$$\begin{cases} x + y &= 1 \\ x - y &= 2 \\ 2x - y &= 3 \end{cases}.$$

We write the augmented matrix and perform ERO on it:

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & -1 & 3 \end{array}\right) \xrightarrow{R_2 - R_1} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 2 & -1 & 3 \end{array}\right) \xrightarrow{R_3 - 2R_1} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & -3 & 1 \end{array}\right)$$

$$\xrightarrow[-\frac{1}{3}R_3]{-\frac{1}{2}R_2} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & -1/2 \\ 0 & 1 & -1/3 \end{array}\right) \xrightarrow{R_3 - R_2} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1/6 \end{array}\right).$$

The last row means that $0 \cdot x + 0 \cdot y = 1/6$, which is impossible. Thus the system has *no* solutions at all.