Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-09-25

Darij Grinberg

December 5, 2019

1. Introduction to matrices (continued)

1.1. Properties of matrix operations (continued)

Recall: Multiplication of matrices is associative but not commutative.

So products of matrices don't need parentheses (i.e., for example, $ABCDE$ makes sense without putting parentheses in, as long as the little products AB, BC, CD, DE make sense), but the order of the matrices in the product matters (so $AB \neq BA$ in general).

One more thing that does not work for matrices: Back in the set of real numbers, we can cancel (i.e., if $a \neq 0$ and $ab = ac$, then $b = c$).

For matrices, this is no longer true: It is easy to come up with an example where A is not all zeroes, and $AB = AC$, but nevertheless $B \neq C$.

1.2. The zero matrix

This section follows [lina, §2.10].

Definition 1.2.1. Let $n, m \in \mathbb{N}$. Then, the $n \times m$ **zero matrix** means the matrix $(0)_{1 \leq i \leq n,\ 1 \leq j \leq m}$. This is the $n \times m$-matrix whose all entries are 0. We denote it by $0_{n \times m}$ or, if ambiguity cannot arise, by 0.

For example, the 2×3 zero matrix is $0_{2\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

The following rules are easy to verify:

Proposition 1.2.2. Let $n, m \in \mathbb{N}$.

(a) We have $0_{n\times m} + A = A + 0_{n\times m} = A$ for each $n \times m$-matrix A.

(b) We have $0_{n\times m} A = 0_{n\times p}$ for each $p \in \mathbb{N}$ and each $m \times p$-matrix A.

(c) We have $A 0_{n\times m} = 0_{p\times m}$ for each $p \in \mathbb{N}$ and each $p \times n$-matrix A.

(d) We have $0A = 0_{n\times m}$ for each $n \times m$-matrix A.
(e) We have $\lambda 0_{n\times m} = 0_{n\times m}$ for each number λ.

1.3. The identity matrix

This section follows [lina, §2.11].

Definition 1.3.1. A **square matrix** means an $n \times n$-matrix for some $n \in \mathbb{N}$.

For example, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a square matrix, but $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ is not.

Definition 1.3.2. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix.
The **diagonal entries** of A are the n entries $A_{1,1}, A_{2,2}, \ldots, A_{n,n}$.
All the other $n^2 - n$ entries of A are called **off-diagonal entries**.

For example, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has diagonal entries a, d and off-diagonal entries b, c.

Definition 1.3.3. If i and j are any two objects (e.g., numbers), then we set

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j. \end{cases}$$

This is called the **Kronecker delta**. For example, $\delta_{2,3} = 0$ but $\delta_{3,3} = 1$.

Definition 1.3.4. Let $n \in \mathbb{N}$. Then, the $n \times n$ **identity matrix** means the matrix

$$\left(\delta_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}.$$

This is the $n \times n$-matrix whose diagonal entries are all equal to 1 and whose off-diagonal entries are all equal to 0. It is denoted by I_n or sometimes by I. (Some call it E.)

For example,

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The following proposition is not hard to prove (see [lina, §2.12] for a proof of part **(b)**):

Proposition 1.3.5. Let $n, m \in \mathbb{N}$.
(a) We have $I_n A = A$ for each $n \times m$-matrix A.
(b) We have $A I_m = A$ for each $n \times m$-matrix A.

1.3.1. Powers of a matrix

This section follows [lina, §2.13].

Recall how powers of numbers are defined:

$$a^0 = 1; \qquad a^1 = a; \qquad a^2 = aa; \qquad a^3 = aaa; \qquad a^4 = aaaa; \qquad \ldots.$$

Likewise, we can define powers of a square matrix:

$$A^0 = I_n; \qquad A^1 = A; \qquad A^2 = AA; \qquad A^3 = AAA; \qquad A^4 = AAAA; \qquad \ldots$$

for any $n \times n$-matrix A.

(Note: Powers are only defined for square matrices.)

Example 1.3.6. • If $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, then

$$\begin{aligned}
A^2 &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}; \\
A^3 &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & 1 \end{pmatrix}; \\
&\ldots; \\
A^n &= \begin{pmatrix} 2^n & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{for each } n \in \mathbb{N}.
\end{aligned}$$

• If $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, then

$$\begin{aligned}
A^2 &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}; \\
A^3 &= A^2 A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}; \\
A^4 &= \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}; \qquad A^5 = \begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix}; \qquad \ldots; \\
A^n &= \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \qquad \text{for each } n \in \mathbb{N}.
\end{aligned}$$

Remark 1.3.7. If A is a square matrix, then $A^u A^v = A^{u+v} = A^v A^u$.

Thus, while not any two matrices commute (i.e., $XY = YX$ does not always hold), we still have a lot of situations where matrices do commute: $XY = YX$ always holds if $X = A^u$ and $Y = A^v$.

1.4. Products and transposes

Proposition 1.4.1. Let A be an $n \times m$-matrix and B be an $m \times p$-matrix. Then,

$$(AB)^T = B^T A^T.$$

For a detailed proof, see [lina, Proposition 3.18 **(e)**].

Example 1.4.2. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$. Then,

$$\begin{aligned} AB &= \begin{pmatrix} ax+bz & bw+ay \\ cx+dz & dw+cy \end{pmatrix}, && \text{so} \\ (AB)^T &= \begin{pmatrix} ax+bz & cx+dz \\ bw+ay & dw+cy \end{pmatrix}, && \text{but} \\ B^T A^T &= \begin{pmatrix} ax+bz & cx+dz \\ bw+ay & dw+cy \end{pmatrix}. \end{aligned}$$

Note: Any **row** of A corresponds to a **column** of A^T. More precisely, for any j, we have

$$(\operatorname{row}_j A)^T = \operatorname{col}_j \left(A^T\right).$$

Similarly,

$$(\operatorname{col}_j A)^T = \operatorname{row}_j \left(A^T\right).$$

Hence, transposing a matrix switches the roles of rows and columns.

2. Gaussian elimination

From now on, we mostly follow [Strickland].

2.0.1. Reminder: solving systems of linear equations

Example 2.0.1. Consider the following system of 3 equations in 3 unknowns x, y, z:

$$\begin{cases} 3x + 6y - z = 2 \\ 7x + 4y - 3z = 3 \\ -y + 8z = 1 \end{cases}.$$

How do we solve it?

There is a method called **Gaussian elimination**. In its most systematic form, it proceeds as follows:

- Try to get the variable x to only appear in **one** equation. To do so, we subtract appropriate multiples of one equation that contains x from all the other equations that contain x.
- After this is done, play the same game with the other variables.

First of all, why can we just subtract equations from each other?
Imagine we are solving a system

$$\begin{cases} A(x, y, z) = 0 \\ B(x, y, z) = 0 \\ C(x, y, z) = 0 \end{cases}.$$

Now, if we subtract the second equation from the first, we get

$$\begin{cases} A(x, y, z) = 0 \\ B(x, y, z) - A(x, y, z) = 0 \\ C(x, y, z) = 0 \end{cases}.$$

I claim that this new system has **exactly the same solutions** as the old system.
Indeed:

- Every solution of $\begin{cases} A(x, y, z) = 0 \\ B(x, y, z) = 0 \\ C(x, y, z) = 0 \end{cases}$ is a solution of $\begin{cases} A(x, y, z) = 0 \\ B(x, y, z) - A(x, y, z) = 0 \\ C(x, y, z) = 0 \end{cases}$, because it satisfies $\underbrace{B(x, y, z)}_{=0} - \underbrace{A(x, y, z)}_{=0} = 0 - 0 = 0$.
- Conversely, every solution of $\begin{cases} A(x, y, z) = 0 \\ B(x, y, z) - A(x, y, z) = 0 \\ C(x, y, z) = 0 \end{cases}$ is a solution of $\begin{cases} A(x, y, z) = 0 \\ B(x, y, z) = 0 \\ C(x, y, z) = 0 \end{cases}$, because it satisfies
$$B(x, y, z) = \underbrace{(B(x, y, z) - A(x, y, z))}_{=0} + \underbrace{A(x, y, z)}_{=0} = 0 + 0 = 0.$$

Thus, when we subtract an equation from another, the system's solutions don't change.

The same holds if we multiply an equation by a nonzero number.

The same holds if we swap two equations.

So we have three transformations we can do with a system of equations without changing its set of solutions:

- We can subtract an equation from another.
- We can multiply an equation by a nonzero number.
- We can swap two equations.

Let's try to use these transformations to simplify our system:

$$\begin{cases} 3x + 6y - z = 2 \\ 7x + 4y - 3z = 3 \\ -y + 8z = 1 \end{cases}$$

$$\xrightarrow{\text{multiply eq. 1 by } 1/3} \begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ 7x + 4y - 3z = 3 \\ -y + 8z = 1 \end{cases}$$

$$\xrightarrow{\text{multiply eq. 2 by } 1/7} \begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ x + \frac{4}{7}y - \frac{3}{7}z = \frac{3}{7} \\ -y + 8z = 1 \end{cases}$$

$$\xrightarrow{\text{subtract eq. 1 from eq. 2}} \begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ -\frac{10}{7}y - \frac{2}{21}z = -\frac{5}{21} \\ -y + 8z = 1 \end{cases} .$$

Now, the x only survives in the first equation.

Let us try to make the y only survive in the second equation:

$$\begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ -\frac{10}{7}y - \frac{2}{21}z = -\frac{5}{21} \\ -y + 8z = 1 \end{cases}$$

$$\overset{\text{multiply eq. 2 by } -7/10}{\longrightarrow} \begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ y + \frac{1}{15}z = \frac{1}{6} \\ -y + 8z = 1 \end{cases}$$

$$\overset{\text{multiply eq. 3 by } -1}{\longrightarrow} \begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ y + \frac{1}{15}z = \frac{1}{6} \\ y - 8z = -1 \end{cases}$$

$$\overset{\text{subtract eq. 2 from eq. 3}}{\longrightarrow} \begin{cases} x + 2y - \frac{1}{3}z = \frac{2}{3} \\ y + \frac{1}{15}z = \frac{1}{6} \\ -\frac{121}{15}z = -\frac{7}{6} \end{cases} .$$

At this point, we can solve eq. 3 for z, getting

$$z = \left(-\frac{7}{6}\right) / \left(-\frac{121}{15}\right) = \frac{35}{242}.$$

With this value found, we can solve eq. 2 for y, getting $y = \frac{1}{6} - \frac{1}{15} \cdot \frac{35}{242} = \frac{19}{121}$. Finally, with this value found as well, we can solve eq. 1 for x, getting $x = \frac{2}{3} - 2 \cdot \frac{19}{121} + \frac{1}{3} \cdot \frac{35}{242} = \frac{97}{242}$. So the only solution to our system is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{97}{242} \\ \frac{19}{121} \\ \frac{97}{242} \end{pmatrix}.$$

Thus our method was successful: We have found a solution of our system, and we know that it is the only solution.

Not all situations are this simple!

Example 2.0.2. Consider the system

$$\begin{cases} x - y = 2 \\ x - z = 3 \\ y - z = 4 \end{cases} .$$

Use the same method:

$$\begin{cases} x - y = 2 \\ x - z = 3 \\ y - z = 4 \end{cases}$$

$$\xrightarrow{\text{subtract eq. 1 from eq. 2}} \begin{cases} x - y = 2 \\ y - z = 1 \\ y - z = 4 \end{cases}$$

$$\xrightarrow{\text{subtract eq. 2 from eq. 3}} \begin{cases} x - y = 2 \\ y - z = 1 \\ 0 = 3 \end{cases} .$$

So this system has no solutions, because 0 will not equal to 3 no matter what x, y, z are.

Example 2.0.3. Consider the system

$$\begin{cases} x - y = 2 \\ x - z = 6 \\ y - z = 4 \end{cases} .$$

Use the same method:

$$\begin{cases} x - y = 2 \\ x - z = 6 \\ y - z = 4 \end{cases}$$

$$\xrightarrow{\text{subtract eq. 1 from eq. 2}} \begin{cases} x - y = 2 \\ y - z = 4 \\ y - z = 4 \end{cases}$$

$$\xrightarrow{\text{subtract eq. 2 from eq. 3}} \begin{cases} x - y = 2 \\ y - z = 1 \\ 0 = 0 \end{cases}$$

$$\xrightarrow{\text{remove eq. 3, since it always holds}} \begin{cases} x - y = 2 \\ y - z = 1 \end{cases} .$$

How to solve $\begin{cases} x - y = 2 \\ y - z = 1 \end{cases}$?

Take z arbitrary, and solve eq. 2 for y, getting $y = z + 1$. With this value found, solve eq. 1 for x, getting $x = \underbrace{y}_{=z+1} + 2 = z + 3$. So the solutions of our system are all triples $(x, y, z) = (z + 3, z + 1, z)$. In other words, they can be written as $(z, z, z) + (3, 1, 0)$.

So we have solved three systems of equations, and seen three different behaviors:

- A unique solution.
- No solution.
- Infinitely many solutions.

So it appears that the answer to an arbitrary system of linear equations will be: Either there will be no solutions at all, or $\underbrace{\text{some}}_{\text{possibly none}}$ variables will be free and others will be determined by them.

2.1. Linear equations vs. matrices

This section follows [Strickland, §4].

To get a better grip on solving system of linear equations, we are going to restate them "in matrix form".

Example 2.1.1. Start again with our system

$$\begin{cases} 3x + 6y - z = 2 \\ 7x + 4y - 3z = 3 \\ -y + 8z = 1 \end{cases} \qquad \text{in three variables } x, y, z.$$

I claim that it is equivalent to the matrix equation

$$\begin{pmatrix} 3 & 6 & -1 \\ 7 & 4 & -3 \\ 0 & -1 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

Indeed,

$$\begin{pmatrix} 3 & 6 & -1 \\ 7 & 4 & -3 \\ 0 & -1 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x + 6y + (-1)\, z \\ 7x + 4y + (-3)\, z \\ 0x + (-1)\, y + 8z \end{pmatrix} = \begin{pmatrix} 3x + 6y - z \\ 7x + 4y - 3z \\ -y + 8z \end{pmatrix}.$$

For this vector to be equal to $\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$, we need precisely the equations

$$\begin{cases} 3x + 6y - z = 2 \\ 7x + 4y - 3z = 3 \\ -y + 8z = 1 \end{cases} .$$

But this is exactly our system that we started with.

Example 2.1.2. The system

$$\begin{cases} x - y = 2 \\ x - z = 6 \\ y - z = 4 \end{cases}$$

rewrites as

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}.$$

This form has some advantages: Now we are solving one equation instead of three, and furthermore we are using matrices, which simplifies our life if we know anything about matrices.

Upshot: For any system of n linear equations in m variables $x_1, x_2, \ldots, x_m$, we can rewrite the system as a matrix equation

$$Ax = b, \qquad \text{where } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Namely, if our equations are

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m = b_2 \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m = b_n \end{cases} ,$$

then the system is equivalent to $Ax = b$, where

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Therefore, solving a system of linear equations is the same as "unmultiplying" a vector by a matrix: We are looking for all x's with $Ax = b$. In a way, we are thus looking for "b divided by A" (except that this is not a thing).

In our above "method" for solving systems, we have used three operations:

- Subtracting an equation from another.
- Multiplying (i.e., scaling) an equation by some $\lambda \neq 0$.
- Swapping two equations.

What do these operations correspond to in terms of the matrix A and the vector b?

- Subtracting a row of A from another row of A, and simultaneously doing the same for b.
- Scaling a row of A by some $\lambda \neq 0$, and doing the same for b.
- Swapping two rows of A, and doing the same for b.

Example 2.1.3. Let us solve the system $\begin{cases} x - y = 2 \\ x - z = 6 \\ y - z = 4 \end{cases}$ using this "method":

$$\begin{cases} x - y = 2 \\ x - z = 6 \\ y - z = 4 \end{cases} \qquad \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}$$

$$\overset{\text{subtract eq. 1 from eq. 2}}{\longrightarrow} \begin{cases} x - y = 2 \\ y - z = 4 \\ y - z = 4 \end{cases} \qquad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 4 \end{pmatrix}$$

$$\overset{\text{subtract eq. 2 from eq. 3}}{\longrightarrow} \begin{cases} x - y = 2 \\ y - z = 1 \\ 0 = 0 \end{cases} \qquad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}.$$

So our operations on the system of equations can be recast as operations on A and b.

2.2. The augmented matrix

This section follows [Strickland, §4].

The easiest way to do some operations on A and b simultaneously is to glue the vector b to the matrix A (let's say on its right edge) and just perform these operations on the resulting big matrix. This is called the **augmented matrix** of our system of equations:

Definition 2.2.1. Let

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m = b_2 \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m = b_n \end{cases}$$

be a system of n linear equations in m unknowns $x_1, x_2, \ldots, x_m$. Then, the matrix

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} & b_n \end{pmatrix}$$

is called the **augmented matrix** of the system. Often you put a vertical bar between the last column and the other columns in this matrix, so as to remind yourself that they serve different roles. I will instead put the entries in the last column in boldface:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} & \mathbf{b}_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} & \mathbf{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} & \mathbf{b}_n \end{pmatrix}.$$

Now, the operations we did on our equations (subtracting, scaling and swapping) correspond directly to operations done on the rows of this augmented matrix. As we already saw, the solutions of the system of equations do not change when we do these operations.

Question: How simple can we make our augmented matrix just by performing these operations?

We need to specify a notion of "simple" matrices. The one we will use is called a "reduced row-echelon matrix" or a "matrix in reduced row-echelon form (RREF)" or "RREF matrix".

2.3. Reduced row-echelon form

This section follows [Strickland, §5].

Definition 2.3.1. Let A be a matrix. We say that A is in **reduced row-echelon form (RREF)** if the following conditions hold:

- **RREF0:** Any row of zeros is at the bottom of the matrix, after all the nonzero rows. (A "nonzero row" means a row that is not all zeros. It can still have some zeros.)

- **RREF1:** In any nonzero row, the first nonzero entry is equal to 1. This entry is called the **pivot** of the row.
- **RREF2:** The pivot of any nonzero row must be further to the right than the pivot of the previous nonzero row.
- **RREF3:** If a **column** contains a pivot, then all other entries in the column are zero.

Example 2.3.2. Here are some examples of matrices that are in RREF and matrices that aren't:

- If an $n \times n$-matrix is in RREF and has n pivots, how does it look like?

  It has to be the identity matrix.

  For example, if $n = 4$, then it has to be $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.
- Is the matrix $\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 5 \end{pmatrix}$ in RREF? Yes.
- Is the matrix $\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ in RREF? No, as it fails both RREF2 and RREF3.
- Is the matrix $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ in RREF? No, because it fails RREF2 (the pivot in row 3 is not right of the pivot in row 2).
- Is the matrix $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ in RREF? No, because it fails RREF0.

What is the purpose of RREF matrices? Idea:

1. We can transform any matrix into an RREF form by the row operations we discussed (subtracting a row from another; scaling a row; swapping two rows).
2. If the augmented matrix of a system of linear equations is in RREF, then this system is very easy to solve.

Combined, these two points give a method for solving systems of linear equations:

Form the augmented matrix; transform it into RREF; then solve the resulting easy system.

Example for point 2: For example, let's say the augmented matrix of our system is

$$\left(\begin{array}{cccc} 1 & 0 & 0 & \mathbf{3} \\ 0 & 0 & 1 & \mathbf{5} \end{array}\right).$$

This means that our system is

$$\begin{cases} 1x + 0y + 0z = 3 \\ 0x + 0y + 1z = 5 \end{cases}, \qquad \text{i.e.}$$

$$\begin{cases} x = 3 \\ z = 5 \end{cases}.$$

This is already the solution. (Here, y can be arbitrary.)

Slightly less trivial example: Let's say the augmented matrix is

$$\left(\begin{array}{cccc} 1 & 4 & 0 & \mathbf{3} \\ 0 & 0 & 1 & \mathbf{5} \end{array}\right).$$

This means that our system is

$$\begin{cases} 1x + 4y + 0z = 3 \\ 0x + 0y + 1z = 5 \end{cases}, \qquad \text{i.e.}$$

$$\begin{cases} x + 4y = 3 \\ z = 5 \end{cases}.$$

The solutions of this system are trivial to find: z has to be 5 (in order to satisfy $z = 5$); then, y can be arbitrary; then, x has to be $3 - 4y$ (in order to satisfy $x + 4y = 3$).

References

[lina] Darij Grinberg, *Notes on linear algebra*, version of 13 December 2016.
`https://github.com/darijgr/lina`

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
`http://neil-strickland.staff.shef.ac.uk/courses/MAS201/`