Mathematical preliminaries of crypto
Linear algebra

Sven Laur

May 10, 2017

1 What are matrices and what are they good for?

In the high school you probably have studied how to solve linear equations and something about matrices. In particular, there were formulae how to solve linear equations of two variables of type:

$$\begin{cases} ax + by = c \\ dx + ey = f \end{cases} \tag{1}$$

where $a, \cdots, f$ are fixed constants and x and y are unknown variables. Often, it is more convenient to use matrix notation to write down the same equation:

$$\begin{pmatrix} a & b \\ e & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ f \end{pmatrix} \ ,$$

as the latter decouples variables and their coefficient. In particular, we can manipulate equation systems without writing down variables. For instance, a step by step solution

$$\begin{cases} 1x - 2y = 2 \\ 2x + 4y = 12 \end{cases} \Rightarrow \begin{cases} 1x - 2y = 2 \\ 0x + 8y = 8 \end{cases} \Rightarrow \begin{cases} 1x - 2y = 2 \\ 0x + 1y = 1 \end{cases} \Rightarrow \begin{cases} 1x - 0y = 4 \\ 0x + 1y = 1 \end{cases}$$

can be written in terms of matrices as follows:

$$\begin{pmatrix} 1 & -2 & 2 \\ 2 & 4 & 12 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -2 & 2 \\ 0 & 8 & 8 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix} \ .$$

More formally, let A be an $m \times n$ matrix with elements a_{ij} and B be an $n \times \ell$ matrix with elements b_{jk}. Then their matrix product $C = AB$ is an $m \times \ell$ matrix with entries

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} \ .$$

In other words, to compute c_{ik} we have to take the ith row from A and write it on top of kth column in B, multiply elements in the same column and then sum the results:

Row	a_{i1}	a_{i2}	$\cdots$	a_{in}	
Column	b_{1j}	b_{2j}	$\cdots$	b_{nj}	Sum
Products	$a_{i1}b_{1j}$	$a_{i2}b_{2j}$	$\cdots$	$a_{in}b_{1j}$	c_{ij}

Matrices with dimensions $1 \times n$ and $n \times 1$ are special. These are called *row* and *column vectors*, respectively. These are commonly used to store values corresponding to a single object or observation. For instance, we can measure students in the class and store their height, weight and age as a three element vectors. However, such vectors are not very meaningful, since their elements are in different dimensions (meters vs kilogrammes vs years) and the sum of elements is not very meaningful. Things get more cumbersome if vector elements are of different type, e.g. the first element is an integer, the second element is residue modulo 5 and the third element is element from a finite field $\mathbb{F}_{16}$. In such case, the sum of all elements is not well defined.

To avoid all kind of complications mathematicians are have agreed that all vector and matrix elements must be of same type that supports addition, subtraction, multiplication, division and follows standard arithmetic laws. That is, all matrix and vector elements are the elements of a same field. In most everyday applications, the underlying field is the set real numbers, whereas finite fields $\mathbb{Z}_p$ and $\mathbb{F}_q$ are more common in cryptography.

Now given two matrices of same dimensions we can compute elementwise sum:

$$\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1b} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m2} & \cdots & a_{mn}+b_{mb} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1b} \\ \vdots & \ddots & \vdots \\ b_{m2} & \cdots & b_{mb} \end{pmatrix}$$

or multiply matrix elements with a same coefficient c:

$$\begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix} = c \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Often, it is also convenient to use transposition to flip the matrix over main diagonal. As a result $m \times n$ matrix becomes $n \times m$ matrix

$$\begin{pmatrix} a_{11} & a_{21} \cdots & ca_{1m} & \\ a_{12} & a_{22} \cdots & ca_{1m} & \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & c_{2n} \cdots & ca_{mn} & \end{pmatrix} = \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}^T$$

These three natural definitions allow us to use constructs of type $A+B$, $c \cdot A$ and A^T to simplify some mathematical derivations of some complex formulae.

Exercise 1. *Compute the expression* $5 \cdot A + BC$ *over real numbers and over* $\mathbb{Z}_2$ *where*

$$A = \begin{pmatrix} 1 & 5 \\ 4 & 3 \\ 5 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad c = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}^T .$$

Exercise 2. *A priory it is not clear whether addition and multiplication for matrices follows standard arithmetic laws. Prove the following facts:*

1. $(A^T)^T = A$;
2. $A + B = B + A$;
3. $A + (-1)A = 0$ *where* 0 *denotes a matrix consisting of zeroes*
4. $c(A + B) = cA + cB$;
5. $(A + B)\boldsymbol{x} = A\boldsymbol{x} + B\boldsymbol{x}$ *for a column vector* $\boldsymbol{x}$;
6. $(A + B)C = AC + BC$ *for a matrix* C *by using the previous fact;*
7. $(\boldsymbol{xy})^T = \boldsymbol{y}^T\boldsymbol{x}$ *for a row vector* $\boldsymbol{x}$ *and a column vector* $\boldsymbol{y}$;
8. $(AB)^T = B^T A^T$ *by using the previous fact.*

2 Linear systems and Gaussian elimination algorithm

Let us now consider a system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \qquad\qquad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

which can be compactly represented in the matrix form as

$$A\boldsymbol{x} = \boldsymbol{b} \ .$$

The algorithm behind equation solving is on tree observations. First, if we add two equations we get an new valid equation:

$$\begin{cases} \alpha_1 x_1 + \cdots + \alpha_n x_n = \beta \\ \gamma_1 x_1 + \cdots + \gamma_n x_n = \delta \end{cases} \qquad \Rightarrow \qquad (\alpha_1 + \gamma_1)x_1 + \cdots + (\alpha_n + \gamma_n)x_n = \beta + \delta$$

Second, we can always multiple equations with some constant

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta \qquad \Rightarrow \qquad \gamma\alpha_1 x_1 + \cdots + \gamma\alpha_n x_n = \gamma\beta \ .$$

Third, we can reorder the equations for better clarity.

Gaussian elimination is a systematic use of these observations to solve the equation. As the first step, we get rid of equations by collecting all coefficients into a single matrix

$$(A|\boldsymbol{b}) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right) \ .$$

Note that all three ways to derive new equations can be modelled with row operations. In particular, we can add a multiple of a one row to another row in the matrix; we can multiply matrix rows by coefficients; and we can switch rows of matrices. Additionally, we can drop rows consisting only from zeroes, as they pose no restrictions to variables.

The first task in Gaussian elimination is to convert the matrix into one of the upper-triangular forms:

$$\begin{pmatrix} * & * & \cdots & * & \cdots & * & * \\ 0 & * & \cdots & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & * & \cdots & * & * \end{pmatrix} \quad \begin{pmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & * & * \end{pmatrix} \quad \begin{pmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & * \end{pmatrix}$$

where $*$ denote arbitrary elements. To which form the matrix reduces shows whether we can solve the system or not. If the system can be converted to the leftmost form then there can be many or no solutions. If the system can be converted to the rightmost form there are no solutions. The middle form with non-zero elements on the diagonal indicates that there is unique solution to the initial system of equations.

To reach this form, we must zero the all elements in the first column except the topmost entry by adding multiples of a row to another row. For instance, consider a system of linear equations

$$\begin{cases} x_1 + 2x_2 = 3 \\ 2x_1 + x_2 = 4 \\ 2x_1 + 3x_2 = 3 \end{cases}.$$

Then we can first multiply the first row by two and then subtract the result form the second row:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 2 & 3 & 3 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 2-2\cdot 1 & 1-2\cdot 2 & 4-2\cdot 3 \\ 2 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 2 & 3 & 3 \end{pmatrix}.$$

Alternatively, we could have subtracted the second row from the third

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 2 & 3 & 3 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 0 & 2 & -1 \end{pmatrix}.$$

In fact, this is more promising as we can then zero the second row

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 2 & 3 & 3 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 0 & 2 & -1 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 2 & -1 \end{pmatrix}.$$

As a result, we have derived new equations

$$\begin{cases} -3x_2 = -2 \\ 2x_2 = -1 \end{cases}.$$

The next goals is to zero all elements in the second column except the top two elements. In case of our example, we can do the following transformations:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 2 & -1 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 1 & -\frac{1}{2} \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -\frac{1}{2} \\ 0 & -3 & -2 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} \end{pmatrix} .$$

Note that the last row corresponds to the unsatisfiable equation

$$0 = -\frac{1}{2}$$

and thus the original system of linear equations is not solvable.

When the first step in Gaussian elimination yields the middle form then we must continue with row operations until we reach a diagonal form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_n \end{pmatrix}$$

Note that this matrix corresponds to system of linear equations

$$\begin{cases} x_1 = c_1 \\ \cdots \\ x_n = c_n \end{cases}$$

and thus the last column $c_1, \ldots c_n$ is the solution to original system.

The strategy to achieve this goal consists of two sub-steps. First, we must multiply each rows by the inverse element of the diagonal entry. If some diagonal element is zero then the system has many solutions and we cannot continue.

Exercise 3. *Find a solution to the equation $A\boldsymbol{x} = \boldsymbol{b}$ when*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$$

and $\boldsymbol{b} = (1\ 1\ 1)^T$ over $\mathbb{Z}_2$, $\mathbb{Z}_3$ and $\mathbb{Z}_5$.

3 What to do if Gaussian elimination fails?

Some errors are irrecoverable. Namely, equation of type $0 = c$ cannot have any solution if c is non-zero coefficient. Thus, we cannot to anything when the Gaussian elimination produces a row of type

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & * \end{pmatrix} .$$

However, we could still proceed if some of the diagonal elements are zeros. For that note that reordering of matrix columns just means reordering of variables. For example, consider the case

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} 1x_1 + 0x_2 + 1x_3 = 1 \\ 0x_1 + 0x_2 + 1x_3 = 0 \\ 0x_1 + 0x_2 + 1x_3 = 1 \end{cases}.$$

Now if we change the location of the second and the third column the second place will encode x_3 and the third place will encode x_2:

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} 1x_1 + 1x_3 + 0x_2 = 1 \\ 0x_1 + 1x_3 + 0x_2 = 0 \\ 0x_1 + 1x_3 + 0x_2 = 1 \end{cases}.$$

Hence, we can swap matrix column provided that we keep track what column encodes what variable. As a result, we can always transform the original matrix into one of the forms

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_m \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & \cdots & 0 & * & \cdots & c_1 \\ 0 & 1 & \cdots & 0 & * & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & * & \cdots & c_m \end{pmatrix}$$

or discover a contradiction. The leftmost form has a unique solution. The rightmost form has many solutions. In particular, the assignment

$$\begin{cases} x_1 = c_1 \\ \cdots \\ x_m = c_m \\ x_{m+1} = 0 \\ \cdots \\ x_n = 0 \end{cases}$$

is a satisfying solution. However, there can be many other solutions. Note that the equations encoded by the matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & * & \cdots & c_1 \\ 0 & 1 & \cdots & 0 & * & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & * & \cdots & c_m \end{pmatrix}$$

can be rewritten in the form

$$\begin{pmatrix} x_1 \\ \cdots \\ x_m \end{pmatrix} + B \begin{pmatrix} x_{m+1} \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \cdots \\ c_m \end{pmatrix}$$

where B contains all $*$ elements. Hence, $x_{m+1}, \ldots, x_n$ can have arbitrary values as long as

$$\begin{pmatrix} x_1 \\ \cdots \\ x_m \end{pmatrix} = \begin{pmatrix} c_1 \\ \cdots \\ c_m \end{pmatrix} - B \begin{pmatrix} x_{m+1} \\ \cdots \\ x_n \end{pmatrix} \quad .$$

Exercise 4. *Find all solutions of the equation $A\boldsymbol{x} = \boldsymbol{b}$ where*

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

and $\boldsymbol{b} = (1\ 2\ 4)^T$ over $\mathbb{Z}_3$

4 How to solve many linear equations simultaneously?

Consider the case when you want to solve linear equations $A\boldsymbol{x} = \boldsymbol{b}$ for many values of $\boldsymbol{b}$. Then you can use the following property

Lemma 1. *If $A\boldsymbol{x}_1 = \boldsymbol{b}_1$ and $A\boldsymbol{x}_2 = \boldsymbol{b}_2$ then $A(\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2) = \alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b_2}$.*

Proof. Indeed, note $A(\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2) = \alpha_1 A\boldsymbol{x}_1 + \alpha_2 A\boldsymbol{x}_2 = \alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b_2}$. □

To make our task easier, we can first find solution for special vectors

$$\begin{aligned} \boldsymbol{e}_1 &= (1\ 0\ \cdots\ 0)^T \\ \boldsymbol{e}_2 &= (0\ 1\ \cdots\ 0)^T \\ &\cdots \\ \boldsymbol{e}_m &= (0\ 0\ \cdots\ 1)^T \end{aligned}$$

which have only a single non-zero component. Let $\boldsymbol{x}_1, \ldots \boldsymbol{x}_m$ be the corresponding solutions, i.e., $A\boldsymbol{x}_i = \boldsymbol{e}_i$. Note that we can express

$$\boldsymbol{b} = b_1\boldsymbol{e}_1 + \cdots + b_m\boldsymbol{e_m}$$

and Lemma 1 assures that

$$A(b_1\boldsymbol{x}_1 + \cdots b_m\boldsymbol{x}_m) = b_1\boldsymbol{e}_1 + \cdots + b_m\boldsymbol{e}_m = \boldsymbol{b} \quad .$$

Moreover, let $X = (\boldsymbol{x}_1|\boldsymbol{x}_2|\cdots|\boldsymbol{x}_m)$ be a matrix with column vectors $\boldsymbol{x}_i$ then $\boldsymbol{Xb}$ is the solution.

Exercise 5. *Prove this claim by showing that $X\boldsymbol{b} = b_1\boldsymbol{x}_1 + \cdots b_m\boldsymbol{x}_m$.*

Exercise 6. *Find the corresponding matrix X over $\mathbb{Z}_7$ for the matrices*

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 5 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 5 & 4 & 6 \end{pmatrix} .$$

For $n \times n$ square matrices A, the solution matrix X is either unique or we cannot find solution for some equation $A\boldsymbol{x} = \boldsymbol{e}_i$. If this matrix X exists then it is called inverse of A since

$$AX = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I$$

and this matrix I does not change matrices if it is multiplied with another matrix. That is, it acts as the element 1 in the fields.