Lecture 2: 1.4 Matrix product. An $m\times n$ matrix is a collection of $m\cdot n$ numbers

$$\begin{array}{ccc} & j \text{ th } \textbf{column} & \\ A = & \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ & & \cdot & & \\ & & \cdot & & \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ & & \cdot & & \\ & & \cdot & & \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} & i \text{ th } \textbf{row} \end{array}$$

The matrix multiplication of an $m \times n$ matrix A by an $n \times 1$ column vector $\mathbf{x}$ is

$$A\mathbf{x} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_m,$$

where $\mathbf{a}_j$ denotes the j-th column of $A = [\mathbf{a}_1 \cdots \mathbf{a}_j \cdots \mathbf{a}_n]$.
Multiplication by an $m \times n$ matrix A defines a linear map $\mathbf{R}^n \ni \mathbf{y} \to A\mathbf{y} \in \mathbf{R}^m$:

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}, \qquad A(\alpha\mathbf{x}) = \alpha A\mathbf{x}.$$

Suppose that A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then multiplication by B defines a map $\mathbf{R}^p \ni \mathbf{x} \to B\mathbf{x} \in \mathbf{R}^n$ and multiplication by A defines a map $\mathbf{R}^n \ni \mathbf{y} \to A\mathbf{y} \in \mathbf{R}^m$ so multiplication by first B and then A

$$\mathbf{x} \xrightarrow{\text{multiply by } B} B\mathbf{x} \xrightarrow{\text{multiply by } A} A(B\mathbf{x})$$

defines a map $\mathbf{R}^p \ni \mathbf{x} \to A(B\mathbf{x}) \in \mathbf{R}^m$. We want to define the matrix product AB to be the $m \times p$ matrix that represents this map so that $(AB)\mathbf{x} = A(B\mathbf{x})$:

$$\mathbf{x} \xrightarrow{\qquad\text{multiply by } AB\qquad} (AB)\mathbf{x} = A(B\mathbf{x})$$

Another way to formulate this is to say that we want the matrix multiplication to be defined so it is **associative**, i.e. $(AB)\mathbf{x} = A(B\mathbf{x})$, it shouldn't matter if you first calculate $B\mathbf{x}$ and then $A(B\mathbf{x})$ or if you first calculate AB and then $(AB)\mathbf{x}$.
Question: How can we find the matrix for the composition of linear maps if we know the matrices for the maps themselves? Let us calculate $A(B\mathbf{x})$ and see

$$B\mathbf{x} = [\,\mathbf{b}_1 \ \cdots \ \mathbf{b}_p\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p.$$

and hence by linearity

$$A(B\mathbf{x}) = A\big(x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p\big) = x_1 A\mathbf{b}_1 + \cdots + x_p A\mathbf{b}_p = [\,A\mathbf{b}_1 \ \cdots \ A\mathbf{b}_p\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

can be interpreted as the matrix product of the $m \times p$ matrix with columns $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$ and the column vector $\mathbf{x}$. Since we already know how to calculate $A\mathbf{b}_j$ where $\mathbf{b}_j$ is a column vector this allows us to define the **matrix multiplication** to be

$$AB = [\, A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p \,]$$

and we have achieved that $(AB)\mathbf{x} = A(B\mathbf{x})$. (That the columns of the matrix of the transformation $\mathbf{x} \to A(B\mathbf{x})$ are $A(B\mathbf{e}_j) = A\mathbf{b}_j$ also follows from section 1.9.)

For practical calculations by hand it is more efficient to use the alternative **row-column rule** to compute the (i, j)th entry of AB as the dot product between the ith row of $A = [\, a_{ij} \,]$ and jth column of $B = [\, b_{ij} \,]$:

$$(AB)_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$$

$$i\text{ th }\textbf{row}\quad \begin{bmatrix} \\ a_{i1}\cdots a_{in} \\ \\ \end{bmatrix} \underset{j\text{ th }\textbf{column}}{\begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix}} = \underset{j\text{ th }\textbf{column}}{\begin{bmatrix} \vdots \\ \cdots (AB)_{ij} \cdots \\ \vdots \end{bmatrix}} \quad i\text{ th }\textbf{row}$$

Quest When is the product AB of an $m \times n$ matrix A and a $q \times p$ matrix B defined?
Question Is matrix multiplication **commutative**, i.e. is $AB = BA$?
Why do people expect things to be commutative in math when they are not commutative in real life? It is not the same thing to first put on the shoes and then the socks as it is to first put on the socks and then the shoes?
What if A is a 2×3 and B is 3×2? Are AB and BA defined?

We skipped the example below per popular request.

Ex Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$. Find AB

Sol $$A\mathbf{b}_1 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-1) + 2 \cdot 1 \\ 0 \cdot (-1) + (-1) \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

$$A\mathbf{b}_2 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 \\ 0 \cdot 1 + (-1) \cdot 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Hence

$$AB = A\Big[\mathbf{b}_1 \mathbf{b}_2\Big] = \Big[A\mathbf{b}_1 A\mathbf{b}_2\Big] = \begin{bmatrix} 1 & 5 \\ -1 & -2 \end{bmatrix}$$

Alternatively using the row-column method

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1(-1) + 2 \cdot 1 & 1 \cdot 1 + 2 \cdot 2 \\ 0(-1) + (-1)1 & 0 \cdot 1 + (-1)2 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -1 & -2 \end{bmatrix}$$

Alternatively one can also write this as

$$\begin{array}{cc} & \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix} \\ AB = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 5 \\ -1 & -2 \end{bmatrix} \end{array}$$

1.5 Triangular factors and row exchange. Let us solve the system in Ex 6 Lecture 1 again:

$$\begin{aligned} x_1 - 2x_2 + \ \ x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

or written in terms of multiplication by the coefficient matrix A

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

If we add 4 times the first row to the last we get the equivalent system

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 0 & -3 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix} \qquad (3) + 4(1)$$

Adding 4 times the first row to the last is equivalent to multiplying both sides with an elementary matrix E

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ 4r_1 + r_3 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 0 & -3 & 13 \end{bmatrix}$$

Dividing the second row by 2:

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -4 \\ 0 & -3 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -9 \end{bmatrix} \qquad (3)/2$$

is equivalent to multiplying with the matrix F

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2/2 \\ r_3 \end{bmatrix}$$

Finally, adding 3 times the first row to the last

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix} \qquad (3) + 3(1)$$

is equivalent to multiplying with the elementary matrix G

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 + 3r_1 \end{bmatrix}$$

We have hence show that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$$

Since the inverse operation to adding a row is subtracting it we conclude that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and since the inverse operation to diving a row by 2 is multiplying it by 2 we conclude that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying by inverses it hence follows that

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$$

The multiplication of lower triangular matrices is a lower triangular matrix:

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ -7 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$$

We have thus obtained the so called *LU* factorization of the matrix into a lower triangular matrix L multiplied by an upper triangular matrix U.

The *LU* factorization doesn't always work directly as is seen by the following examples

$$\begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}$$

In this example one first have to multiply with a permutation matrix P interchanging the rows:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}$$

However one can prove that one can always factorize any matrix $A = PLU$, where P is a permutation matrix.