Theory sheet 1

Definition 1. *A matrix is a rectangular array of numbers:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Here m is the number of rows and n is the number of columns. Size of the matrix is a pair of numbers, written as $m \times n$. The set of $m \times n$ matrices is denoted by $M_{m,n}(\mathbb{R})$ or just $M_{m,n}$.

- *A matrix is said to be a column vector if the numer of columns is one: $n = 1$.*
- *A matrix is said to be a row vector if the number of rows is one: $m = 1$.*
- *A matrix is said to be square if $m = n$. Any non-square matrix is said to be rectangular.*
- *A matrix is said to be symmetric if $(A)_{ij} = (A)_{ji}$.*
- *A matrix is said to be diagonal if $(A)_{ij} = 0$ for all $i \neq j$.*
- *Zero matrix is denoted 0 or $0_{m\times n}$ if we want to specify its size. All its elements are zero.*
- *Identity matrix is denoted I or I_n, it is defined by $(I)_{ij} = 1$ if $i = j$ and $(I)_{ij} = 0$ otherwise. Note that I is a square matrix.*

Definition 2. *Matrix addition and multiplication by a scalar is definied componentwise:*

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij}, \quad (\lambda A)_{ij} = \lambda (A)_{ij}.$$

Note that the sum of two matrices of different sizes is undefined.

Definition 3. *The product of two matrices $A \in M_{m,p}$ and $B \in M_{p,n}$ is a new matrix $AB \in M_{m,n}$ with elements defined by*

$$(AB)_{ij} = \sum_{k=1}^{p} (A)_{ik}(B)_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Note that

- The product AB is only defined if $A \in M_{m,p}$ and $B \in M_{p,n}$. If the number of columns in A is different from the number of rows in B, the product is undefined.
- The product AB is not defined componentwise: $(AB)_{ij} \neq (A)_{ij}(B)_{ij}$.

- $AB = 0$ does not imply that $A = 0$ or $B = 0$. The simplest examples are
$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \times 1 + (-1) \times 1 = 0.$$
It is said that there are divisors of zero in $M_{m,n}$: pairs of non-zero matrices with zero product.

- Matrix product is non-commutative: $AB \neq BA$.

- Due to non-commutativity, many identities such as $(x-y)(x+y) = x^2 - y^2$ are not valid in the space of matrices:
$$(A - B)(A + B) = A^2 - AB + BA - B^2.$$
The part in blue is not in general zero. Similarly, the identity $(x+y)^2 = x^2 + 2xy + y^2$ is not valid:
$$(A + B)^2 = A^2 + AB + BA + B^2.$$
The part in blue is not in general equal to $2AB$.

Definition 4. *The column space of a matrix A is the set*
$$L_c := \left\{ \boldsymbol{x} = \alpha_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \alpha_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + \alpha_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \quad \alpha_1, \ldots, \alpha_n \in \mathbb{R} \right\}.$$
Note that L_c is a vector space: if $\boldsymbol{x}, \boldsymbol{y} \in L_c$ and $\lambda \in \mathbb{R}$, then $\boldsymbol{x} + \boldsymbol{y} \in L_c$ and $\lambda \boldsymbol{x} \in L_c$.

Definition 5. *The row space of a matrix A is the set*
$$L_r := \left\{ \begin{aligned} &\boldsymbol{x} = \beta_1 \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{pmatrix} + \beta_2 \begin{pmatrix} a_{21} & a_{22} & \ldots & a_{2n} \end{pmatrix} \\ &\quad + \cdots + \beta_m \begin{pmatrix} a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \quad \beta_1, \ldots, \beta_m \in \mathbb{R} \end{aligned} \right\}.$$
Note that L_r is also a vector space: if $\boldsymbol{x}, \boldsymbol{y} \in L_r$ and $\lambda \in \mathbb{R}$, then $\boldsymbol{x} + \boldsymbol{y} \in L_r$ and $\lambda \boldsymbol{x} \in L_r$.

Definition 6. *Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be linearly independent if $\alpha \boldsymbol{x} + \beta \boldsymbol{y} = \boldsymbol{0}$ implies $\alpha = \beta = 0$. A collection of vectors $\boldsymbol{x}_i$, $i = 1, \ldots, n$ is linearly independent if $\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_n \boldsymbol{x}_n = \boldsymbol{0}$ implies that $\alpha_1 = \cdots = \alpha_n = 0$.*

- Example: vectors $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ are linearly independent.

- More generally, vectors
$$\begin{pmatrix} a_{11} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} * \\ a_{22} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} * \\ * \\ a_{33} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} * \\ * \\ * \\ a_{44} \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} * \\ * \\ * \\ * \\ \vdots \\ a_{mn} \end{pmatrix}$$
are linearly independent if $a_{11}, a_{22}, a_{33}, a_{44}, \ldots, a_{mn} \neq 0$.

- Counterexample: vectors $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}3\\0\end{pmatrix}$ are linealy dependent.

Definition 7. *Dimension of a vector space L is the maximal number of linearly independent vectors in this space. It is denoted by $\dim L$.*

- *The only space with $\dim L = 0$ is $L = \{\mathbf{0}\}$.*
- *If L is a subspace of $\mathbb{R}^n$, then $\dim L \leq n$. For example, any subpsace of a plane has dimension less or equal than 2.*

The following important theorem is given without proof:

Theorem 1. *Dimensions of the column space and of the row space of a matrix are equal.*

This theorem justifies the following definition:

Definition 8. *Rank of a matrix is the dimension of its column space or the dimension of its row space:*

$$\operatorname{Rank} A = \dim L_r = \dim L_c.$$

- *Since $\dim L_r \leq n$ and $\dim L_c \leq m$, then $\operatorname{Rank} A \leq \min\{m, n\}$. This is important for doing sanity checks: if you have found the rank of a 2×3 matrix to be equal 3, there is a mistake in your arguments because it should be no bigger than 2!*
- *It follows from the remark in the definition of linear independence that if A is of the form*

$$A = \begin{pmatrix} a_{11} & * & * & * & \cdots & * & * & \cdots \\ 0 & a_{22} & * & * & \cdots & * & * & \cdots \\ 0 & 0 & a_{33} & * & \cdots & * & * & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{pn} & * & * & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

with some $a_{11}, a_{22}, \ldots, a_{pn} \neq 0$, then $\operatorname{Rank} A = p$ because the first p columns are linearly independent. This statement is an important tool for finding $\operatorname{Rank} A$ as we shall see now.

Definition 9. *Elementary transformations on rows are the following three operations defined on a matrix:*

- *Adding one row to another. For example,*

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} + a_{21} & a_{12} + a_{22} & \cdots & a_{1n} + a_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

We denote such transformations by $L_j \leftarrow L_j + L_i$, which reads "replace j^{th} line by the sum of j^{th} line and the i^{th} line".

- *Multiplying a row by a non-zero constant λ. For example,*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

  This transformation is usually denoted by $L_i \leftarrow \lambda L_i$.

- *Swapping rows. For example,*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

  This transformation is usually denoted by $L_i \leftrightarrow L_j$.

Definition 10. *Elementary operations on columns are defined similarly.*

Elementary transformations on rows and columns are important because of the following easy theorem (try proving it!):

Theorem 2. *Elementary operations on rows do not change the row space of a matrix. Elementary operations on columns do not change the column space of a matrix.*

The following corollary immediately follows from the last theorem:

Corollary 1. *Elementary transformations do not change rank of a matrix.*

This allows us to calculate the rank of a matrix by playing the following game:

Transform a given matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

to the form for which finding the rank is easy:

$$A' = \begin{pmatrix} a_{11} & * & * & * & \cdots & * & * & \cdots \\ 0 & a_{22} & * & * & \cdots & * & * & \cdots \\ 0 & 0 & a_{33} & * & \cdots & * & * & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \dots \\ 0 & 0 & 0 & \dots & a_{pn} & * & * & \cdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots \end{pmatrix} \Longrightarrow \text{Rank } A' = p$$

by applying elementary transformations. By corollary above, $\text{Rank } A = \text{Rank } A' = p$.

Remark 1. *Mixing row and column elementary transformations is fine as long as as we are calculating the rank. We will, however, use elementary transformations for several other things. In these other problems one must not mix these and stick to only one kind of transformations. To avoid confusion from now on we will only use row operations.*

If we decide to avoid column operations, it may not be possible to bring a matrix to the form

$$\begin{pmatrix} a_{11} & * & * & * & \cdots & * & * & \cdots \\ 0 & a_{22} & * & * & \cdots & * & * & \cdots \\ 0 & 0 & a_{33} & * & \cdots & * & * & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & \dots & a_{pn} & * & * & \cdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots \end{pmatrix}.$$

For example, to transform the following matrix

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

to the said form, we would need to swap columns. Fortunately, finding rank of such matrices is easy without swapping columns: just note that the highlighted columns

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

are linearly independent. Since rank is the maximal number of linearly independent columns, we see that the rank of this matrix is 2 *. In other words, to find rank it is sufficient to bring the matrix to a form*

$$\begin{pmatrix} 0 & \cdots & 0 & a_{11} & * & \cdots & * & * & \cdots & * & * & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & a_{22} & * & \cdots & * & * & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & a_{33} & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots \end{pmatrix}$$

and just count the number of linearly independent columns:

$$\begin{pmatrix} 0 & \cdots & 0 & a_{11} & * & \cdots & * & * & \cdots & * & * & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & a_{22} & * & \cdots & * & * & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & a_{33} & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots \end{pmatrix}.$$