Lecture 2: Linear Algebra

Advanced Microeconomics I, ITAM

Xinyang Wang*

In this section, we explore the linear structure of $\mathbb{R}^n$. The purpose of this section is to introduce a basic understanding of the linear structure of a set, as well as some basic matrix operations. The linear structure sets will play a key role in the next section on the convexity of sets and functions, while the matrix operations will be useful in both multivariate calculus (comparative statics) and optimizations.

1 Linear Spaces

We start our discussion by the linear space $\mathbb{R}^N$.[1]

*Please email me at xinyang.wang@itam.mx for typos or mistakes. Version: January 17, 2021.

[1] A linear space (also called a vector space) consists of a non-empty set V together with addition and multiplication by numbers, denoted by $v + w$ and cv, where $v, w \in V$ and $c \in \mathbb{R}$ such that the following ten property holds:

1. for all $v, w \in V$, $v + w \in V$,
2. for all $v \in V$, $c \in \mathbb{R}$, $cv \in V$,
3. for all $v, w \in V$, $v + w = w + v$,
4. there exists a $0 \in V$ such that $0 + v = v$ for all $v \in V$,
5. for all $v \in V$, there exists a unique $-v \in V$ such that $v + (-v) = 0$,
6. for all $v \in V$, $1v = v$
7. for all $v \in V$, $c_1, c_2 \in \mathbb{R}$, $(c_1c_2)v = c_1(c_2v)$,
8. for all $v, w, z \in V$, $(v + w) + z = v + (w + z)$,
9. for all $v, w \in V$, $c \in \mathbb{R}$, $c(v + w) = cv + cw$,
10. for all $v \in V$, $c_1, c_2 \in \mathbb{R}$, $(c_1 + c_2)v = c_1v + c_2v$.

It is straightforward to check $V = \mathbb{R}^N$ with the traditional definition of $+, \times$ satisfies all these properties.

1.1 Subspace

Now, we define the linear subspaces of $\mathbb{R}^N$:
we say a non-empty set $X \subset \mathbb{R}^N$ is a *subspace of* $\mathbb{R}^N$, or simply a subspace, if

- for all $x, y \in X$, $x + y \in X$
- for all $x \in X$, $c \in \mathbb{R}$, $cx \in X$

Sometimes, we call these two conditions by that X is closed under addition and scalar multiplication.

The following exercise might be helpful to develop a geometric sense about the subspaces: it is good to draw the following sets in $\mathbb{R}^2$ and determine whether they are subspaces of $\mathbb{R}^2$ or not.

Exercise. *Determine whether the following sets are subspaces of* $\mathbb{R}^2$*, draw these sets on a plane, and for the sets that are not subspaces, use pictures to illustrate why they are not subspaces of the plane.*

- $\{(x, x) : x \in \mathbb{R}\}$
- $\{(x, -x) : x \in \mathbb{R}\}$
- $\{(x, x) : x \in \mathbb{R}\} \cup \{(x, -x) : x \in \mathbb{R}\}$
- $\{(x, x) : x \in [-1, 1]\}$

Proposition. *Some basic properties of subspaces are given as follows.*

1. $\{0\}, \mathbb{R}^N$ *are subspaces of* $\mathbb{R}^N$
2. 0 *is in any subspace of* $\mathbb{R}^N$
3. *the intersection of subspaces is a subspace.*
4. *the union of subspaces may not be a subspace.*

Exercise. *Prove the above proposition.*

By repeatedly applying the definition of a subspace, we know if $x_1, ..., x_n$ is in a subspace, then so is $c_1x_1 + ... + c_nx_n$ for any numbers $c_1, ..., c_n \in \mathbb{R}$. Intuitively, in $\mathbb{R}^2$, we could use a singleton $\{(1,1)\}$ to represent the subspace $\{(x,x) : x \in \mathbb{R}\}$, and a set $\{(1,1),(1,-1)\}$ to "represent" the subspace $\mathbb{R}^2$, in the sense that a smallest subspaces of $\mathbb{R}^2$ containing these sets must be the subspace we specified.

Now, we ask whether the above observation can be generalize. That is, for any subspace of $\mathbb{R}^N$, whether or not there is a finite subset such that it represents the subspace. To answer this question, we introduce the concepts linearly independence and basis.

1.2 Linear Independence

Given vectors $x_1, ..., x_n$ in $\mathbb{R}^N$, we say the set $\{x_1, ..., x_n\}$ is *linearly independent* if the equation

$$\alpha_1 x_1 + ... + \alpha_n x_n = 0, \alpha_1, ..., \alpha_n \in \mathbb{R} \tag{1}$$

has a unique solution $\alpha_1 = ... = \alpha_n = 0$. Otherwise, we say $\{x_1, ..., x_n\}$ are *linearly dependent*.

Remark.

1. *We note equation (1) always has a zero solution. Thus, the set is linearly independent if and only if equation (1) has a nonzero solution.*

2. *When the set is linearly dependent, the nonzero solution of equation (1) may not be all nonzero. For instance, $\{(1,0,0),(0,1,0),(1,1,0),(0,0,1)\}$ is a linearly dependent set in $\mathbb{R}^3$. But the solution of equation (1) is $\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = -1, \alpha_4 = 0$.*

Example. *We give some examples of linearly independent and linearly dependent sets. Here x_1, x_2, x_3 are nonzero vectors in $\mathbb{R}^N$, c is some nonzero number.*

- $\{(1,0,0,...,0),(0,1,0,...,0),...,(0,...,0,1)\}$ *is linearly independent*
- $\{x_1\}$ *is linearly independent*
- $\{x_1, 0\}$ *is linearly dependent*
- $\{x_1, cx_1\}$ *is linearly dependent*

- $\{x_1, x_2, x_1 + x_2\}$ *is linearly dependent*
- $\{x_1, x_2, x_3, x_1 + x_2 + x_3\}$ *is linearly dependent*

We say a vector $x \in \mathbb{R}^N$ is a *linear combination* of $\{x_1, ..., x_n\}$ if $x = \alpha_1 x_1 + ... + \alpha_n x_n$ for some numbers $\alpha_1, ..., \alpha_n \in \mathbb{R}$.

It is clear that the concept linear combination is related to linear dependence. The following exercise might be helpful to distinguish them:

Exercise. *Determine whether the following statement is true or false. If it is true, prove it. If it is false, give a counter example: If a set $\{x_1, ..., x_n\} \subset \mathbb{R}^N$ is linearly dependent, then x_1 is a linear combination of $\{x_2, ..., x_n\}$.*

1.3 Span

Next, we formally define what it means by a set generalize a space:

For a nonempty set $S \subset \mathbb{R}^N$, the *span* of S is defined to be the set of linear combinations of finitely many elements in S. In particular, when S is a finite set in the form $S = \{x_1, ..., x_n\}$, then

$$span(S) = \{\alpha_1 x_1 + ... + \alpha_n x_n : \alpha_1, ..., \alpha_n \in \mathbb{R}\}$$

The span of any set must be a linear subspace of $\mathbb{R}^N$:

Proposition. *For any nonempty set $S \subset \mathbb{R}^N$, $span(S)$ is a subspace of $\mathbb{R}^N$.*

Proof. Exercise. ∎

For this reason, if $span(S) = X$ for some subspace X, we say the set S *generates (or spans)* X, or the space X *is generated by (or spaned by)* S.

1.4 Basis

Finally, we are ready to define a basis of a subspace, which is a smallest set generates the subspace:

Given a subspace X of $\mathbb{R}^N$, if a set $\{x_1, ..., x_n\} \subset \mathbb{R}^N$ is linearly independent and $span\{x_1, ..., x_n\} = X$, we say $\{x_1, ..., x_n\}$ is a *basis* of X, and the dimension of X, $\dim(X)$, is n.

Moreover, when a subspace X of $\mathbb{R}^N$ has dimension $N-1$, we say X is a *hyperplane*.

Now, we give some examples to illustrate the above definition:

Example.

- $\{(1,0),(0,1)\}$ *is a basis of* $\mathbb{R}^2$
- $\{(1,1),(-1,1)\}$ *is a basis of* $\mathbb{R}^2$
- $\{(1,0),(1,1)\}$ *is a basis of* $\mathbb{R}^2$. *Hence a basis may not be unique.*
- $\{e_1, ..., e_n\}$ *is a basis of* $\mathcal{R}^n$*, where* $e_k = (0, .., 0, 1, 0, ...0)$ *is the vector having an* 1 *at its* k*-th coordinate and* 0*s everywhere else.*[2]
- $span\{(1,0,0),(0,1,0)\}$ *is a hyperplane in* $\mathbb{R}^3$. *(Draw the hyperplane.)*
- $span\{(1,1,1)\}$ *is not hyperplane in* $\mathbb{R}^3$.*(Draw the set.)*

So far, we defined a set satisfying certain properties is a basis of a subspace. However, we have not yet known whether a basis of a subspace always exists. Therefore, we need to prove the following existence theorem:

Theorem. *Any subspace has a basis.*

One proof of this theorem needs two lemmas, both of which are helpful observations by themselves.

Proposition 1. *In a* n*-dimensional subspace* X*, any linearly independent set in* X *has a cardinality no larger than* n.

Proposition 2. *(Basis Completion Lemma) If* $\{x_1, ..., x_n\}$ *is a linearly independent in* X*. Then either we have* $span\{x_1, ..., x_n\} = X$*, or there is an* $x_{n+1} \in X$ *such that* $\{x_1, .., x_{n+1}\}$ *is linearly independent.*

In the following, we first prove the theorem, then hint the proof of these two propositions in order.

[2] This basis is called *the standard basis of* $\mathbb{R}^n$.

Proof. (Proof of the theorem) We prove the theorem by construction. First, we pick any nonzero vector in X, name it as x_1. Clearly, $\{x_1\}$ is linearly independent. Then we repeatedly apply the basis completion lemma. By Proposition 2, either $\{x_1\}$ spans the whole space X, in which case we are done, or we can find an x_2 such that $\{x_1, x_2\}$ is linearly independent in X. Repeat this process, and either we have a linearly independent set spans X, or we have a longer linearly independent set in X. However, by proposition 1, this process of enlarging linear independent set can not go on forever. Therefore, at some point, we have a linearly independent set generates X. That is, a basis of X. ■

Remark. *To construct a basis of* $\mathbb{R}^3$*, we start with a vector* $x_1 = (1, 0, 0)$*. Since span*$\{(1, 0, 0)\}$*, which is the x-axis, is not the whole space, we can find another vector* x_2 *such that* $\{(1, 0, 0), x_2\}$ *is linearly independent. Say the vector is* $(1, 1, 0)$*. Now, we know span*$\{(1, 0, 0), (1, 1, 0)\}$*, which is the x-y plane in* $\mathbb{R}^3$*, is not the whole space. Therefore, we can find* x_3*, say is* $(1, 1, 1)$ *such that* $\{(1, 0, 0), (1, 1, 0), x_3\}$ *is linearly independent. Then, we have a basis of* $\mathbb{R}^3$*. It is clear that such basis is not unique.*

Hint for proving Proposition 1. We first notice that an equivalent way to state Proposition 1 is that any homogeneous system with n linear equations[3] and more than n variables must have a non-zero solution. For instance, for a homogeneous system with 1 equation and 2 variables, such as $2x + 3y = 0$, there is a nonzero solution.

In this hint, we will only use the equivalence to prove the proposition. To show that any homogeneous system with n linear equations and more than n variables must have a non-zero solution, one need to be familiar with the Gaussian elimination. See, for instance, Chapter 7.1 in Simon and Blume.

Now, we prove by contradiction. Suppose that in $\mathbb{R}^n$, there are $n+1$ linearly independent vectors $x_1, ..., x_{n+1}$, where

$$x_j = \begin{pmatrix} v_{j,1} \\ \vdots \\ v_{j,n} \end{pmatrix} \in \mathbb{R}^n$$

[3] a system of linear equation is homogeneous if every equation is in the form $a_1x_1 + ... + a_nx_n = 0$. That is, the constant terms in the equations are all 0.

And we study the equation

$$\alpha_1 \begin{pmatrix} v_{1,1} \\ \vdots \\ v_{1,n} \end{pmatrix} + ... + \alpha_{n+1} \begin{pmatrix} v_{n+1,1} \\ \vdots \\ v_{n+1,n} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2}$$

Each row of the above equation will gives us one equation. There are n equations in total, but there are $n+1$ variables $\alpha_1, ..., \alpha_{n+1}$. Therefore, equation (2) has a nonzero solution. Contradiction. ■

Hint for proving Proposition 2. We prove by assuming there are no $x_{n+1} \in X$ such that $\{x_1, ..., x_{n+1}\}$ is linearly independent. Since $\{x_1, ..., x_n\}$ is linearly independent, we know any $x \in X$ is a linear combination of $\{x_1, ..., x_n\}$. (Why?) That is, $span\{x_1, ..., x_n\} = X$. ■

2 Matrix

In this section, we will just focus on some basic operations on matrices. These operations will help us to understand multivariate calculus and optimizations. In addition, we will also very briefly relate the matrix with some additional applications. Interested readers can read the relevant material by themselves.

For natural numbers m, n, an *m by n matrix* A is an m by n array of numbers in the form

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

where a_{ij} are all numbers for any $1 \le i \le m, 1 \le j \le n$. The ij-th entry of matrix A (the entry at the i-th row and j-th column of matrix A) is usually written as A_{ij}. When there are no confusions, I will write such a matrix A by $(a_{ij})_{1\le i\le m, 1\le j\le n}$, or simply (a_{ij}).

Now, we give some examples of matrices

Example.

- *A vector is an m by n matrix where $n = 1$.*

- *A square matrix is an m by n matrix where $m = n$.*

- *A zero matrix is a matrix such that all entries are 0.*

- *An identity matrix is a square matrix I such that its diagonal entries are 1 and all other entries are 0. i.e.*

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

  In matrix multiplications, the identity matrix is like number 1 in number multiplications.

- *A diagonal matrix is a square matrix D such that and all off-diagonal entries are 0. i.e.*

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

- *An upper triangular matrix is a square matrix $A = (a_{ij})$ such that $a_{ij} = 0$ for all $i > j$. i.e.*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2(n-1)} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{(n-1)(n-1)} & a_{(n-1)n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

- *A symmetric matrix is a square matrix $A = (a_{ij})$ such that $a_{ij} = a_{ji}$ for all $1 \leq i, j \leq n$.*

2.1 Matrix Operations

In this subsection, we define the matrix addition, scalar multiplication, transpose and matrix multiplications.

First, we notice that the set of matrices is a subspace of $\mathbb{R}^{mn}$. Here, addition and scalar multiplication means the following:

- For m by n matrices A and B, $A+B$ is an m by n matrix such that

$$(A+B)_{ij} = A_{ij} + B_{ij}, \forall 1 \leq i \leq m, 1 \leq j \leq n$$

- For an m by n matrix A and a number $c \in \mathbb{R}$, cA is an m by n matrix such that

$$(cA)_{ij} = cA_{ij}, \forall 1 \leq i \leq m, 1 \leq j \leq n$$

Next, the *transpose* of an m by n matrix A, denoted by A^T, is an n by m matrix, where

$$(A^T)_{ij} = A_{ji}, \forall 1 \leq i \leq n, 1 \leq j \leq m$$

Last, we define matrix multiplications. For an m by k matrix A and a k by n matrix B, their product AB is an m by n matrix, where

$$(AB)_{ij} = \sum_{l=1}^{k} A_{il} B_{lj}, \forall 1 \leq i \leq m, 1 \leq j \leq n$$

Another way to remember the above relation is that the ij-th entry of AB is the dot product of the i-th row of A and the j-th column of B.

Two important special case of matrix multiplications are given as below:

- For a matrix A and a vector v,

$$Av = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + ... + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + ... + a_{mn}v_n \end{pmatrix}$$

- Given two vectors $v, w \in \mathbb{R}^n$, their *dot product* is given by

$$v \cdot w = v^T w = v_1 w_1 + ... + v_n w_n$$

Some properties of matrix operations are given as follows:

Proposition. *Let A, B, C be square matrices, I be the identity matrix, O be the zero matrix, D_1, D_2 be two diagonal matrices, if all these matrices have the same size, we have*

1. $A + O = O + A = A$
2. $0A = A$
3. $(A^T)^T = A$
4. $AI = IA = A$
5. $AO = OA = O$
6. $(AB)C = A(BC)$
7. $(AB)^T = B^T A^T$
8. *AB may not be equal to BA.*
9. *AA^T may not be equal to $A^T A$.*
10. *$D_1 D_2$ is a diagonal matrix, and is equal to $D_2 D_1$.*
11. *There exists a nonzero n by n matrix A such that $A^{n-1} = O$. When $n = 2$, an example could be*

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

Proof. It is an exercise to verify the above proposition. ■

2.2 Sign of a matrix

For a smooth function $f : \mathbb{R} \to \mathbb{R}$, we recall that f is convex (concave) at a point if and only if $f''(x) \geq (\leq) 0$. That is, the convexity of a real-valued function on $\mathbb{R}$ is determined by the sign of its second derivative. The same observation is true for real-valued functions on $\mathbb{R}^n$. Although in such cases, the second derivative would be a matrix, named Hessian matrix, which will be introduced in a later section. But, here, we are ready to ask the question: how to define the sign of a matrix.

Definition 1. *For an n by n matrix A, we say*

- *it is positive semi-definite, c.f. ≥ 0, if $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$.*
- *it is positive definite, c.f. > 0, if $x^T Ax > 0$ for all nonzero $x \in \mathbb{R}^n$.*
- *it is negative semi-definite, c.f. ≤ 0, if $x^T Ax \leq 0$ for all $x \in \mathbb{R}^n$.*
- *it is negative definite, c.f. < 0, if $x^T Ax < 0$ for all nonzero $x \in \mathbb{R}^n$.*

Remark.

1. $x^T Ax = Ax \cdot x = \sum_{1\leq ij \leq} a_{ij} x_i x_j$.
2. *As we will see later, a function is convex/ strictly convex/ concave/ strictly concave at a point x if and only if its Hessian matrix at x is positive semi-definite/ positive definite/ negative semi-definite/ negative definite.*
3. *A matrix is positive semi-definite/ positive definite/ negative semi-definite/ negative definite if and only if all its eigenvalues[4] are non-negative/ positive/non-positive/negative.*
4. *There are some matrices that are neither positive semi-definite nor negative semi-definite. (See example below)*

Example.

1. $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ *is positive definite.*
2. $A = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$ *is negative definite.*
3. $A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ *is neither positive semi-definite nor negative semi-definite.*

The corresponding functions of these three examples are $f(x, y) = \vec{x}^T A\vec{x}$, where $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$. That is, the Hessian matrix of function f defined as this is given by the three matrices

[4] A number λ is an *eigenvalue* of A if $Ax = \lambda x$ for some nonzero vector $x \in \mathbb{R}^n$. It is straightforward to check a diagonal matrix, with diagonal entries $d_1, ..., d_m$, has eigenvalues $d_1, ..., d_n$.

above. In other word, f is convex, concave and neither convex nor concave at point $(0,0)$*, respectively. In the language of optimizations,* $(0,0)$ *is a local minimizer, local maximizer, saddle point of the function f, respectively.*

2.3 Other Applications

Other than helping us to study the property of demand in this course, matrices is ubiquitous in both research works and the real world. Here, we give just a few more examples among the numerous applications.

2.3.1 System of Linear Equations

A system of linear equations can also be written in the matrix form as

$$A\vec{x} = b$$

For instance,

$$\begin{cases} 3x + 2y = 5 \\ 2x + 3y = 6 \end{cases} \iff A\vec{x} = b$$

where

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}, b = \begin{pmatrix} 5 \\ 6 \end{pmatrix}, \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

The benefit of writing in the matrix form is that we could convert the problem of solving a system of linear equations to a problem of studying the property of the matrix A. In particular, whether A is invertible, or of full rank. Most numerical methods in optimizations will be reduced to solving some system of linear equation in the form $Ax = b$.

2.3.2 Linear Transformations

We say a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *linear* if

- $f(x) + f(y) = f(x + y)$, for all $x, y \in \mathbb{R}^n$
- $f(cx) = c(fx)$, for all $x \in \mathbb{R}^n, c \in \mathbb{R}$

It could be proved that a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if $f = Ax$ for some m by n matrix A.[5]

2.3.3 Graph Theory

Graph theory has a wide range of applications in pure and applied mathematics, economics, engineering and computer science. Here in this part, we just remark that the study of a graph is closely related to the study of a corresponding matrix.

A *graph* is a pair (V, E), where V is a finite set of vertices and $E \subset \{\{x, y\} : x, y \in V, x \neq y\}$ is a set of edges. For instance, when $V = \{1, 2, 3, 4\}$ and $E = \{\{1, 3\}, \{2, 4\}, \{3, 4\}\}$, we have a graph with 4 vertices, in which the vertices 1 and 3, 2 and 4, 3 and 4 are connected.

Given any graph (V, E), we can define its adjacency matrix A, which is an $|V|$ by $|V|$ matrix such that

$$A_{ij} = \begin{cases} 0, & \text{if i and j are not linked} \\ 1, & \text{if i and j are linked} \end{cases}$$

For the above example, the adjacency matrix is

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

In applications, graphs is usually very large in the sense that there are a large number of vertices. Therefore, the size of this matrix is large. Interesting properties of a graph includes the centrality and the connectivity of a graph. For instance, in the first category, given the connectivity between webpages/ facebook users, we wish to identify the important website/ famous persons. (c.f. The Pagerank Algorithm) In the second category, we wish to study how well connected a graph is. A related problem is that, if assuming there are some fixed network among people, and in a pandemic, the disease can be transmitted from one person to another in this network with some probability. The question that whether a pandemic can be

[5] To prove this observation, we first note it is clear that f(x)=Ax is linear. Conversely, we pick a basis of $\mathbb{R}^n$ and construct A by the function values of the elements in the basis.

controlled eventually is related to the connectivity of this graph. Both of these properties are related to the eigenvalues of an adjacency matrix (or the Laplacian of a graph). It is evident that there are a lot more interesting questions related to networks, including community detection, the Netflix problem (Matrix Completion), the pricing problem of Uber, expender graphs, shortest path problem, and so on.

2.3.4 Big Data Analysis

As we have mentioned in previous sections, real world applications of matrices usually involves solving some equation in the form $A\vec{x} = b$, where A is a very large matrix. In general, solving this equation is impossible (very time consuming) when A is large. However, the computation can be a lot easier if the rank of A is small.[6] Therefore, the study of big data analysis often involves approximating a large matrix by a low rank matrix. For more information on this line, you can search the relation between five factor model of personality and the singular value decomposition.

[6] I skipped the important concept matrix rank in this course. The *rank* of an m by n matrix A is defined to be the dimension of the space $\{Ax : x \in \mathbb{R}^n\}$.