Linear algebra comments

Sophie Marques

Friday 9^{th} October, 2015

Of course this does not cover all the class notes and it is not enough to do the midterm. It is just a way to extract the very very important part of the course and I do not mean that you do not have to know the remaining part.

Main Advise: Do not try to learn by heart, try first to understand what is going on and you will remember.

What have we learnt?

We have learnt at this point how to solve a system of equations, linear transformation, matrices and connection between all these notions.

0.1 Systems

We first wanted to solve a system of linear equation:

$$\left\{\begin{array}{rcl} a_{1,1}x_1 + \cdots + a_{1,n}x_n & = & b_1 \\ \cdots & & \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n & = & b_m \end{array}\right. \qquad (*)$$

To this system we associate two matrix:

1. THE COEFFICIENT MATRIX:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \cdots & \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

2. THE AUGMENTED MATRIX:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} & b_1 \\ & \cdots & & \\ a_{m,1} & \cdots & a_{m,n} & b_m \end{pmatrix}$$

Let $b = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix} \in \mathbb{R}^n$.

The faster way to solve a system is to ROW REDUCE the augmented matrix corresponding to it until you obtain the ROW REDUCE ECHELON FORM.

Write the row reduction that you are doing and make sure it is one of these operation you are doing:

1. Replacement: $R_i \leftarrow R_i + \lambda R_j$, λ a scalar;

2. Scale: $R_i \leftarrow \lambda R_i$, λ a non-zero scalar;

3. Switch: $R_i \leftrightarrow R_j$.

(Remember that we are allowed to do such operation because they lead to equivalent augmented matrix corresponding to equivalent system, so you are not changing the solution set of your system, and that is what you want to find. So, mentioning with the symbol $\sim$ that two matrices are equivalent is important.)
We have seen two other equivalent ways to see the system above:

1. matrix equation form $Ax = b$.

2. vector equation form $a_1x_1 + \cdots + a_nx_n = b$ where a_i are the column of the coefficient matrix A.

So if I ask you solve the system $(*)$ or $Ax = b$ or $a_1x_1 + \cdots + a_nx_n = b$, the method will be the same as describe above unless mentioned differently.
Once you solve a system, you might have to describe the full solution set. If the last column of the augmented matrix is not a pivot column the solution set is empty, and you have nothing to do. Otherwise, the system will have only one solution if there is no free variable here it is also easy or infinitely many if there are free variable and in which case we will express the basics variable in term of the free variables.
There are several way to do this, but most likely you might have to give me the solution set in vector parametric form where the parameter are the free variable. That mean that you give a solution of the system on the form

$$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = v_0 + t_1v_1 + \cdots + t_nv_n$$

where t_i is the parameters and v_i are vector in $\mathbb{R}^n$.
If I ask you for a geometric description. I am looking for how the solution looks like in $\mathbb{R}^n$: If you obtain:

1. $\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = v_0 + t_1 v_1$, this is a line passing through the point of coordinate v_0 with direction v_1.
2. $\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = v_0 + t_1 v_1 + t_2 v_2$, this is a plane passing through the point of coordinate v_0 and with direction v_1 and v_2.
3. $\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = v_0 + t_1 v_1 + t_2 v_2 + t_3 v_3$, this is a space of 3 dimension though the point v_0 and with direction v_1, v_2, v_3

Bigger than this you cannot represent it in your head unfortunately but you can imagine what is going on...

0.2 Vectors

0.2.1 Linear independence

A set of vector $\{v_1, \cdots, v_n\}$ (for instance the set of the column vector of a matrix) in $\mathbb{R}^m$ is linearly independent if the homogeneous vector equation $x_1 v_1 + \cdots + x_n v_n = 0$ has only one solution, the trivial solution. Otherwise we say that the set is linearly dependent.
In order to know if a set of vector is linearly independent of no, you can as usual ROW REDUCE the augmented matrix associated to the system and see if there are free variable of not.
Here a characterization of linear dependence.

Theorem 0.2.1 (Characterization of linearly dependent sets)**.** *An indexed set $S = \{v_1, \cdots, v_r\}$ of two vectors is linearly dependent if and only* *if one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $v_1 \neq 0$ then some v_j (with $j > 1$) is a linear combination of the preceding vectors $v_1, \cdots, v_{j-1}$.*

0.2.2 Span

A set of vector $\{v_1, \cdots, v_n\}$ (for instance the set of the column vector of a matrix) in $\mathbb{R}^m$ spans $\mathbb{R}^m$ if the equation $x_1v_1 + \cdots + x_nv_n = b$ has at least a solution for all $b \in \mathbb{R}^m$. We write also $Span\{v_1, \cdots, v_n\} = \{x_1v_1 + \cdots + x_nv_n, x_1, \cdots, x_n \in \mathbb{R}\} = \mathbb{R}^m$.
Here we have equivalent way to see the spanning problem, and the fourth point will be the most useful in practice since you can just row reduce the matrix and see if you have a pivot position in each row or not. Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

1. For each b in $\mathbb{R}^m$, the equation $Ax = b$ has a solution.
2. Each b in $\mathbb{R}^m$ is a linear combination of the column of A.
3. The columns of A span $\mathbb{R}^m$.
4. A has a pivot position in every row.

0.3 Linear transformation

We have see also linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$. That is a transformation such that

$$T(u + v) = T(u) + T(v),\ \mathit{for\ all}\ u, v \in \mathbb{R}^n$$

and

$$T(cu) = cT(u),\ \mathit{for\ all}\ u \in \mathbb{R}^n\ \mathit{and\ cscalar}$$

You could also just say that

$$T(cu + dv) = cT(u) + dT(v),\ \mathit{for\ all}\ u, v \in \mathbb{R}^n\ \mathit{and}\ c, d\ \mathit{scalars}$$

We also have see that a matrix transformation is a (linear) transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ which maps x into Ax, for a $m \times n$ matrix.

We have prove that ANY linear transformation is a matrix transformation and the corresponding matrix is called a standard matrix. You can compute it with computing the image of the standard vector e_i by T

$$[T(e_1), \cdots, T(e_n)]$$

where e_i is the i^{th} vector column of the identity matrix I_n.
We define the notion of one-to-one and onto mapping:
A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^m$ if each b in $\mathbb{R}^m$ is the image of at least one x in $\mathbb{R}^n$.
A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each b in $\mathbb{R}^m$ is the image of at most one x in $\mathbb{R}^n$.
That is the characterization of one-to-one you will mostly use in practice for linear transformation:

Theorem 0.3.1. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation $T(x) = 0$ has only the trivial solution.*

And now you just have to retranslate in term mentioned before and you will know what to do

Theorem 0.3.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let A be the standard matrix for T. Then:*

1. *T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$ (this is also equivalent to every vector of $\mathbb{R}^m$ is a linear combination of the columns of A) ;*
2. *T is one-to-one if and only if the columns of A are linearly independent.*

0.4 Matrices

1. If A is a $m \times n$ matrix, that is, a matrix with m rows and n columns, then the scalar entry in the ith row and jth column of A is denoted by $a_{i,j}$ and is called the (i, j) entry of A. Each column of A is a list of m real numbers, which identifies with a vector in $\mathbb{R}^m$. Often, these columns are denoted by $a_1, \cdots, a_n$ and the matrix A is written as

$$A = [a_1, \cdots, a_n]$$

Observe that the number $a_{i,j}$ is the ith entry (from the top) of the jth column vector a_j. The **diagonal entries** in an $m \times n$ matrix $A = [a_{i,j}]$ are $a_{1,1}$, $a_{2,2}$, $a_{3,3}$, $\cdots$ and they form the **main diagonal** of A. A **diagonal matrix** is a square $n \times n$ whose non diagonal entries are zero. An example is the $n \times n$ identity matrix. An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0. The size of a zero matrix is actually clear from the context.

2. **(Equality)** Two matrix are **equal** if they have same size (i.e same number or row and same number of column) and if their corresponding columns are equals which amounts to saying that their corresponding entries are equal.
3. **(Sum)** The **sum** of two matrix A and B is defined if and only if the two matrices HAVE THE SAME SIZE. The sum of two matrix $m \times n$ is a matrix $m \times n$ whose columns is the sum of the corresponding columns. That is, the entries of $A+B$ is the sum of the corresponding entries in A and B.
4. **(Scalar multiplication)** If r is a scalar and A is a matrix $m \times n$, then the **scalar multiple** rA is the matrix whose the columns are r times the corresponding column in A. We denote $-A$ for $(-1)A$ and $A - B = A + (-1)B$.
5. [1]et A be a matrix $m \times n$ and B be a matrix $n \times p$ with column b_1, b_2, $\cdots$, b_p, then the **product** AB is defined and it is a $m \times p$ matrix whose column are Ab_1, Ab_2, $\cdots$, Ab_p. That is,

$$AB = A[b_1, \cdots, b_p] = [Ab_1, \cdots, Ab_p]$$

Each column of AB is a linear combination of the column of A using the weights from the corresponding column of B. We have that the (i, j) entry of the matrix AB is

$$(ab)_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

for each $i = 1, \cdots, m$ and $j = 1, \cdots, p$.
Be careful, in order for the multiplication to be defined it is necessary that the NUMBER OF COLUMNS OF A equals the NUMBER

1. L

OF ROWS IN B. Also the matrix AB has size $m \times p$. The number of rows is equal to m (number of row of A) and the number of column is equal to p (number of column of B).

6. **(Powers)** If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of copies of A

$$A^k = A \cdots A$$

If $k = 0$, $A^0 = I_n$.

(WARMINGS)

1. Be careful, in general $AB \neq BA$. We say that A and B **commute** with another if $AB = BA$.DO NOT FORGET THIS IS NOT TRUE IN GENERAL.
2. The cancelation laws DO NOT hold for matrix multiplication. That is $AB = AC$, then it is NOT TRUE in general that $B = C$. You could have $AB = AC$ but still $B \neq C$.
3. If a product AB is the zero matrix, you CANNOT conclude in general that either $A = 0$ or $B = 0$. You could have $AB = 0$ but still $A \neq 0$ or $B \neq 0$.

Given an $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix, denoted by A^T whose columns are formed from the corresponding rows of A. Note that $(AB)^T = B^T A^T$ and $(A^T)^T = A$.

0.4.1 Inverse

INVERTIBLE MATRICES ARE **SQUARE** MATRICES.
An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix denoted A^{-1} such that

$$A^{-1}A = AA^{-1} = I_n$$

where I_n is the $n \times n$ identity matrix. The matrix A^{-1} is called the **inverse** of A.
A matrix that is not invertible is sometimes called a **singular matrix**, and an invertible is called a **non singular** matrix.

Theorem 0.4.1. *Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *If* $ad - bc \neq 0$, *then* A *is invertible and*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

If $ad - bc = 0$, *then* A *is not invertible. The quantity* $ad - bc$ *is called the* **determinant** *of* A, *and we write*

$$det(A) = ad - bc$$

Theorem 0.4.2. *If* A *is invertible* $n \times n$ *matrix, then for each* b *in* $\mathbb{R}^n$, *the equation* $Ax = b$ *has the unique solution* $x = A^{-1}b$.

Note that $(A^T)^{-1} = (A^{-1})^T$, $(AB)^{-1} = B^{-1}A^{-1}$ and $(A^{-1})^{-1} = A$. An **elementary matrix** is one that is obtained by performing a single elementary row operation on the identity matrix.

Theorem 0.4.3. *An* $n \times n$ *matrix* A *is invertible if and only if* A *is row equivalent to* I_n *and in this case. any sequence of elementary row operation that reduces* A *to* I_n *also transform* I_n *into* A^{-1}.

Row reduce the augmented matrix $[A, I]$. If A is row equivalent to I, then $[A, I]$ is row equivalent to $[I, A^{-1}]$, Otherwise, A does not have an inverse.

Theorem 0.4.4 (The invertible matrix theorem)**.** *Let* A *be a square* $n \times n$ *matrix. Then the following statements are equivalent. That is, for a given* A, *the statement are either all true or all false.*

1. *A is an invertible matrix*
2. *A is row equivalent to the $n \times n$ identity matrix.*
3. *A has n pivot positions.*
4. *The equation $Ax = 0$ has only the trivial solution*
5. *The columns of A form a linearly independent set.*
6. *The linear transformation $x \mapsto Ax$ is one-to-one.*
7. *The equation $Ax = b$ has at least one solution for each b in $\mathbb{R}^n$.*
8. *The equation $Ax = b$ has a unique solution for each b in $\mathbb{R}^n$.*

9. *The columns of A span $\mathbb{R}^n$.*
10. *The linear transformation $x \mapsto Ax$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.*
11. *There is an $n \times n$ matrix C such that $CA = I$.*
12. *There is an $n \times n$ matrix D such that $AD = I$.*
13. *A^T is an invertible matrix.*

Now you know, that if you let A and B be square matrices. If $AB = I$, then A and B are both invertible, with $B = A^{-1}$ and $A = B^{-1}$.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S \circ T = T \circ S = I_n$$

That is, for all $x \in \mathbb{R}^n$, $S(T(x)) = x$ and $T(S(x)) = x$.
The next theorem shows that if such an S exists, it is unique and must be a linear transformation. We call S the **inverse** of T and write it as T^{-1}.

Theorem 0.4.5. *Let $T : \mathbb{R} \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case the linear transformation S given by $S(x) = A^{-1}x$ is the unique function satisfying*

$$S \circ T = T \circ S = I_n$$

Note that the standard matrix of the composition of two linear transformation is the product of the standard matrix of this linear transformation (be careful with the order of this product).