MIT Gilbert Strang Linear Algebra Course总结与心得

之前一直听说MIT有全世界最好的线性代数课，听完3节课以后，觉得虽不至于出神入化，但是讲课时知识点的条理性、思路的连贯性、对理解的强调，都是在国内的课上鲜少见到的。

Strang教授讲课还是很可爱的，他的课堂不会很严肃刻板，偶尔抄错一两个数字，黑板空间不够了，还会下意识说一句Damn!

教授上课时说话的风格很像聊天，不会像教科书那样死板。课堂内容的讲解基本上是会用最简单的，幼儿园小朋友也能听懂的单词来表述，而且教授不鼓励同学们在日常学习数学时记忆或者使用太复杂的学术名词。Don’t forget the rule, forget the word “Associative”!

同时，教授在课上经常使用具体例子阐明一种数学思想，而非机械地向学生灌输一个公式或者定理。

目录

[Lec1 – The Geometry of Linear Equations 1](#_Toc124699578)

[Lec2 – Elimination with Matrices 2](#_Toc124699579)

[Lec3 – Multiplication and Inverse Matrices 3](#_Toc124699580)

[Lec4 – Factorization into A=LU 5](#_Toc124699581)

[Lec5 – Transposes, Permutations, Spaces R^n 6](#_Toc124699582)

[Lec6 – Column Space and Nullspace 7](#_Toc124699583)

[Lec7 – Solving Ax=0: Pivot Variables, Special Solutions 8](#_Toc124699584)

[Lec8 – Solving Ax=b: Row Reduced From R 9](#_Toc124699585)

# Lec1 – The Geometry of Linear Equations

关键词：Linear Equations, Row Picture, Column Picture, Matrix form, Vector, Linear Combination

Lec1中，Strang教授强调了线性代数的基本问题：求解线性方程组。这一点似乎国内的课程没有过多强调，但是从前三节课的内容看，理解这一点是极为重要的，因为矩阵的定义与运算规则，本质上都是在求解线性方程组的任务下建立的。

Lec1的重点内容是对一个线性方程组两个角度的理解，即行的角度Row Picture和列的角度Column Picture。Row Picture是常见的做法，以方程为单位，生成每个方程对应的平面的线或空间的面，联立的交点即为方程的解；Column Picture更为重要，体现了线性代数的思想，这种方法将方程系数构造成矩阵，未知数构造成列向量，两者相乘，即将方程转化为列向量的线性组合问题。

同时也提到只有当矩阵是Good Matrix的时候，即可逆，方程才确保一定有解，否则三个列向量的线性组合无法表示三维空间中的任意列向量，则方程有无解和右侧的Right-handed side列向量有关。

# Lec2 – Elimination with Matrices

关键词：Elimination, Triangular Matrix, Elimination Failure, Pivot, Multiplier, Back Substitution, Augmented Matrix, Elimination Matrix, Associative Law, Elementary Matrix, Permutation Matrix

Lec2主要讲了一种直观的方程求法Elimination，以此作为出发点，引出了矩阵重要的数学意义——对方程消元步骤的表示，同时教授强调不会这么快讲行列式。Elimination的思想最早由Gauss提出，通过对方程组系数矩阵行之间的相消，构造Triangular Matrix，从而直观地得出方程的解。最后一步操作称为Back Substitution。

系数矩阵的每一行代表一个方程，对角线上的元素称为Pivot，进行Elimination时用第一行消去第二行的3，用第一、第二行消去第三行的0、4，即得到上三角矩阵。为使等式Right-handed side得以和Left-handed side进行相同的加减消元操作，将Right-handed side列向量补充到的右侧，组成一个Augmented Matrix，同步进行Elimination。

Elimination完成后，原本的系数矩阵成为Triangular Matrix ，所表示的方程组变成如下的等价形式，的值已经显然，依次代入也能直接得出的值。

Elimination Failure的情况是，系数矩阵的对角线Pivot出现0且无法通过换行消除，此时矩阵不是Good Matrix，即不可逆，方程不一定有解。

Lec2最重要的内容是，Expression of Matrices，即利用矩阵相乘代表矩阵变换。这里教授讲了一句很重要的话，体现了矩阵相乘的基本数学意义——行与列的线性叠加，概括就是：矩阵左乘是对行的处理，右乘是对列的处理。

以上矩阵A的Elimination可以用矩阵相乘表示为

是Do it all的矩阵，可以称为Elimination Matrix。由此引出了矩阵的两个重要含义——表示线性方程组和表示矩阵变换步骤。课程亦提到矩阵相乘满足结合律Associative Law，不满足交换律Commutative Law。

将矩阵变换逆过来进行需要用到逆矩阵Inverse Matrix，而逆矩阵与原矩阵相乘则得到单位矩阵Indentity Matrix，这是Do nothing的矩阵，相当于代数乘法的1。

# Lec3 – Multiplication and Inverse Matrices

关键词：Multiplication, Row times Column, Block, Inverses, Invertible, Gauss-Jordan Idea

在这一讲里，教授用了半节课的时间讲矩阵相乘的五种方法，但其实更像是五种不同的理解角度。五种方法分别是行乘列、行叠加、列叠加、列乘行叠加、分块。

第一种方法，行乘列，是最常用的方法，即计算结果的第i行第j列元素Component等于前一个矩阵第i行与后一个矩阵第j列的点积Dot Product，难以体现矩阵相乘的数学意义。

第二种方法，行叠加，体现了行线性组合的思想。将后一个矩阵看作目标矩阵，前一个矩阵看作变换矩阵，且变换矩阵由m行行向量组成，则结果的第i行就是以变换矩阵第i行的行向量各分量作为系数，目标矩阵诸行向量的线性组合。

第三种方法与第二种方法的思想很相似，不过是列组合。后一个是变换矩阵。

行组合和列组合的方法与之前课程上所讲的对线性方程组系数矩阵进行变换的数学意义一脉相承，具有连贯性。

第四种方法是列乘行叠加，这种方法的数学意义尚不理解，但是做法很特别，前一个矩阵n列分别与后一个矩阵的n行一一对应点乘，得出的n个矩阵相叠加就是计算结果。

第五种方法是分块计算的方法，课上这一块讲得不多，只是说分块之间的计算法则很类似正常的矩阵运算。

逆矩阵是这节课另外一个重点，具体内容包括数学意义上的可逆性判断和逆矩阵的一种求法——Gauss-Jordan Idea。从线性方程组的角度理解，一个系数矩阵不可逆的原因是它不能保证任意Right-handed side列向量下方程有解，即矩阵的列向量之间存在线性相关。这种情况下，会有非零的未知数向量使得以下式子成立：

对于可逆矩阵，从Elimination变换的角度理解，逆矩阵是将矩阵所代表的变换过程逆过来进行的矩阵，因此两者相乘时，正变换后逆变换，相当于Do nothing。

Gauss-Jordan Idea从解方程的角度给出了矩阵求逆的一种朴素思想，将未知的逆矩阵分解为多个列向量的线性方程组，同时进行Elimination，这就是矩阵求逆的初等变换法。

# Lec4 – Factorization into A=LU

关键词：Inverses, Transpose, Elimination (Matrix), Factorization, Permutaion Matrix

Lec4的内容比较散乱，大概是作为承上启下的一课。重点内容是A=LU，即将A矩阵的Elimination Matrix E移到等式右侧，可以消除E矩阵中折算的Multiplier，以从L矩阵中看出每一步Elimination直接使用的Multipier。这一课的内容可以听懂，但是不太好理解。

教授首先讲了多个矩阵相乘，再求逆、转置，会改变运算顺序，这里用了一个很生动的比喻：先穿上袜子，再穿鞋；脱鞋，再脱袜子。这一部分是给后面矩阵移位作铺垫。

矩阵求逆、转置可以任意互换顺序：

讲Factorization的时候，先从2x2矩阵讲起，再拓展到3x3矩阵，教授说3x3矩阵Significantly different，因为3x3矩阵的Elimination需要多步，涉及到多个Elimination Matrices，就涉及到矩阵相乘。这节课上的矩阵A默认不需要行变换预处理。

对A进行Elimination的三个矩阵，求逆移到等式右边，组成矩阵L。

举例，假设A进行Elimination的步骤依次是，行一消行二，行一不需要消行三，行二消行三，则有Elimination Matrices：

即使行一对行三没有直接消元，三个矩阵乘起来总的Elimination矩阵中第三行第一列仍然出现10，表示行一通过行二对行三的消元间接影响了行三，这是一个运算顺序引起的折算Multiplier。而这是一个不利因素，导致不能从变换矩阵E中直接看出Elimination每一步的Multiplier。

对E求逆后移到右侧，得到的L矩阵不再有折算Multiplier，因此便可以直观看出Elimination过程中某一行对某一列消元所用的Multiplier（不过符号好像是反过来的），这就是这节课的重要思想。

这个E矩阵变L矩阵的过程不太容易理解，课程也没有给出证明。纯逻辑关系上的解释大概就是一种影响关系的传递。

接下来教授用一些时间讲了Elimination的算法时间复杂度(How expensive?)，最坏情况下，n阶系数矩阵A消元所需时间为

本次课的最后内容是移位矩阵Permutation Matrices，本质上是行或列经过移动的单位矩阵Identity Matrix，乘在矩阵A左侧就对矩阵A进行行移位。对于n阶的矩阵，共有个不同的移位矩阵。移位矩阵的特殊性质是其逆等于转置。

# Lec5 – Transposes, Permutations, Spaces R^n

关键词：Transpose, Permutation, Symmetric Matrix, Vector Space, Subspace

这节课整理了一些矩阵转置、移位、对称矩阵的性质，然后重点以向量的线性组合引出了线性代数的核心内容——线性空间，被教授称为“The beginning of real Linear Algebra”

移位矩阵Permutation Matrices的性质再次被提到，同时指出若需要进行Elimination的A矩阵Pivot的位置有0，可以通过左乘移位矩阵P的方法进行行移位Row Exchange。

矩阵转置Transpose的方法很简单，就是行变成列，列变成行，或者说一个Component的行列编号交换。

对称矩阵Symmetric Matrix的定义是，转置后还是自己。对于任意矩阵R，都是对称矩阵，因为有：

接下来是重点内容——向量空间。向量空间包含所有二维向量，向量空间包含所有n维向量。向量的集合要组成向量空间的条件是满足八大运算规则，但是教授说只需要记住一个词——封闭“Closed”，意即向量空间中的任意向量任意线性组合（加、标量乘）也必须位于空间内。

向量空间的第一象限不构成向量空间，因为这些恒正的向量只要乘一个负标量就会脱离第一象限，不满足封闭性。

向量空间Vector Space的子空间Subspace只包含它一部分的向量，但也能成为一个向量空间，必要条件是必须包含零向量（原点）。对于三维向量空间，其Subspace有以下四种：自身、过原点的面、过原点的线、零向量自身。要注意，过原点的面所代表的向量空间不等价于，因为其本质上仍是三维向量。

由矩阵A的诸列向量生成的向量空间称为A的列空间Column Space of A。

# Lec6 – Column Space and Nullspace

关键词：Column Space, Nullspace, Right-handed side, Solution

这节课开始触及线性代数的核心，从矩阵的列空间、零空间的角度理解线性方程组的可解性，以及解释什么时候有解，为什么有解。

向量空间的其中两个子空间为P和L，容易发现，P和L的并集Union不是的子空间，而交集Intersection是的子空间。

接下来是重头戏。对于矩阵A，它的三个四维列向量生成了一个四维空间中的子空间，这个子空间包含了这三个向量和它们全部的线性组合，称为矩阵A的列空间Column Space，记作C(A)。

线性方程组不一定有解，或者说，它只对特定的Right-handed side向量b有解。从列空间的角度理解，若向量b位于矩阵A的列空间C(A)中，则向量b可以被矩阵A的列向量线性组合表示，方程因而有解。在这个例子中，由于A的列空间C(A)不构成整个向量空间，因此对部分四维向量b来说，方程是无解的。

观察可以发现，矩阵A的三个列向量之间并不线性独立Linear Independent，某一个向量总是可以被另外两个向量线性表示，用教授的话说，就是它“Contribute nothing new”。即使去掉一个向量也不影响Column Space的生成，保留的两个向量称为Pivot。

接下来是零空间，矩阵A的零空间Nullspace指的是所有使方程组成立的解向量x所生成的空间，此时右侧b向量为0。这里的矩阵A零空间N(A)包含以下向量，零空间向量组成中的一条直线。

零空间有其特殊性是因为，在的情况下，方程组的解向量不包含零向量，无法组成一个向量空间。因此方程组的解不能组成向量空间，只有的解可以组成向量空间——零空间。

# Lec7 – Solving Ax=0: Pivot Variables, Special Solutions

关键词：Ax=0, Rank, Null Space, Echelon Form, Reduced Row Echelon Form, Pivot Variables, Free Variables, Special Solutions

这节课主要从向量空间的角度理解Right-handed side b向量为0时的方程组的通解，这些通解组成了称为零空间Nullspace的向量空间。

另外，课上开始讨论非方阵的Rectangular矩阵，以使得出的结论更有普适性。

在进行Elimination解方程的基础上，增加对矩阵的几步操作，以得出两个重要的矩阵形式Echelon Form和Reduced Row Echelon Form，在这两种形式下，我们得以清晰地找出矩阵的Pivot个数，即秩Rank，最终，还能写出方程组的解的一般形式。

下面是课程给出的例子：

例子中，矩阵A的诸行列之间并不是线性独立的，进行Elimination后，得到阶梯状的矩阵U，称为系数矩阵的Echelon Form。“下阶梯”的元素就是矩阵的Pivots。

矩阵的秩Rank就是Pivot的个数，Pivot所在的列是中枢列Pivot Column（列一、列三），这也是矩阵线性独立的列，对应中枢变量Pivot Variables，另外的两行称为自由列Free Column（列二、列四），对应自由变量Free Variables。

由于自由变量对应的列可以由中枢列线性表示，自由变量的值随意改变并不影响解空间——零空间Nullspace的生成。将自由变量分别设为；可以得到两个特解。

自由变量（自由列）的个数决定方程组特解Special Solution的个数，而这些特解的所有线性组合就是方程组的通解All solutions。

为了总结出解的一般形式，接下来将Echelon Form的系数矩阵U进一步化简，通过行变换将Pivot Column除Pivot以外的元素置0，所有Pivot置1，就得到Reduced Row Echelon Form (RREF)。

进行列变换，将Pivot Column移到一起（当然，未知变量的顺序也要作相应变换），可以发现RREF由四个小矩阵块Block组成。

特解组成的零空间矩阵Nullspace Matrix N满足下列方程，因此由RREF可以直接得到的特解。（注意未知量的排列顺序要和对应列的排列顺序一致，Pivot Variables在前，Free Variables在后）

零空间矩阵N的每一列都是的一个特解，因此零空间矩阵的列空间Column Space就是原方程组的零空间，的通解。

# Lec8 – Solving Ax=b: Row Reduced From R

关键词：Ax=b, Solvability, Particular Solution, Rank, Full column, Full row

上节课Lec7提到了方程组的一般形式解，今天继续探讨更一般的形式的有解性以及解的一般形式。这节课有两个重点：方程组的通解、矩阵秩Rank与方程组解的数量的关系。

为了课程连贯性，沿用上节课的系数矩阵作为例子，不过方程右侧b不再为0。

因此，方程Elimination也要带上b，构造增广矩阵Augmented Matrix。

求的分为两步，求特解Particular Solution和求零空间（通解），特解的求法是向自由变量Free Variables任意赋值，求出中枢变量Pivot Variables的值。一个很重要但是不显然的结论是，的通解是其特解加上的通解，以下的式子可以说明。

另方程组的，可以求出一个特解；上节课已经求出了的通解，因此直接可以写出的通解。

值得注意的是，由于方程的通解构成零空间，因此之间的线性组合仍是的解；但是，方程的通解不构成向量空间，因此之间的线性组合不是的解，其通解是零空间平移后的结果。

下面是另一个课程重点——讨论矩阵Rank和方程解的关系。前面的课已经提到，矩阵的秩Rank指的是其Pivot的个数，即线性独立的列数，因此矩阵的秩r不能大于列数n。又因为化为Echelon Form以后一行最多只有一个Pivot，因此秩r不能大于行数m。

列满秩Full Colomn Ranks指的是的情况，只可能发生在的矩阵中。此时矩阵代表的方程所有变量均为Pivot Variables，没有Free Variables，因此的解只有，零空间。此时若方程有解，则仅有唯一解。

行满秩Full Row Ranks指的是在的矩阵中出现，此时系数矩阵的每一行都对应一个Pivot，所有行（方程）线性独立，方程对任意的Right-handed side b有解，且有无穷多个解。

对于方阵而言，则更加特殊，若，直接称满秩Full Ranks，综合了以上两种情况。方阵满秩等价于可逆Invertible，此时方程有且仅有一个解，Echelon Form更是直接变为单位矩阵Identity Matrix

一般情况下，系数矩阵非满秩，其Echelon Form是最一般的形式，方程无解或有无穷多个解。

一句话总结就是，“The rank tells everything about the number of solutions”。