18.06 – Linear Algebra

My Notes

Leonard Mohr

Student at Foothill College

leonardmohr@gmail.com

Contents

Section 1

Lecture 1: The Geometry of Linear Equations

Subsection 1.1

Fundemental Problems of Linear Algebra

1.1.1 Matrix-Vector Multiplication Problem

Definition 1 **The Matrix-Vector Multiplication Problem**

Let $m, n \in \mathbb{N}$. Let $A \in \mathbb{R}^{m\times n}$ be a given matrix and $\mathbf{x} \in \mathbb{R}^n$ be a given vector. Then the matrix-vector multiplication problem is to find an unknown vector $\vec{b} \in \mathbb{R}^m$ such that

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

How does this apply in the real world? Well, computer graphics is an important example. Here it is important to think of a matrix as encoding a linear transformation. Let's say

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}.$$

Note that the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is where $\hat{\imath}$ lands after the transformation and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is where $\hat{\jmath}$ lands after the transfmoration. Thus performing the matrix vector multiplication we can determine where the vector $\vec{x}$ will land after the change in coordinates.

1.1.2 Nonsingular Linear-Systems Problem

One of the fundemental problems of linear equations is solving systems of linear equations. For example let's say we have the equations

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3. \end{aligned}$$

We will often put this equation in matrix form:

$$\underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{b}.$$

Note that A is often called the coefficient matrix, since it is made up of the coeficients of x and y. To "solve" this system of linear equations we find the set of vector(s) X such that $A \cdot X = b$. In this case, that would be the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Definition 2 **Nonsingular Linear-Systems Problem**

Let $n \in \mathbb{N}$. Let $A \in \mathbb{R}^{n\times n}$ be a given nonsingular matrix and $\mathbf{b} \in \mathbb{R}^n$ be a given vector. Then the nonsingular linear-systems problem is to find an unknown vector $\mathbf{x} \in \mathbb{R}^n$ such that

$$A \cdot \mathbf{x} = \mathbf{b}$$

Subsection 1.2

Row Picture

1.2.1 Two Dimensions

Looking at the same linear system of equations as above:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned}$$

giving us the matrix

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3. \end{bmatrix}$$

The row picture is really what we are most familiar with by now; we just plot the equations on our graph, and then where the lines intersect is the solution to the problem.

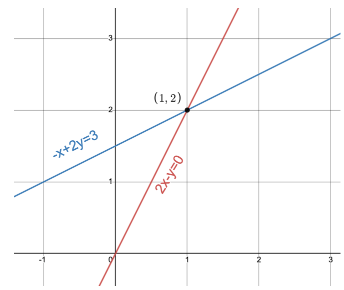


Figure 1. Row picture of our equation $A \cdot X = b$.

1.2.2 Three Dimensions

What if we have three variables and three unknowns:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y - z &= -1 \\ -3y + 4z &= 4. \end{aligned}$$

In matrix form that is

$$\underbrace{\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix}}_{A'} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{X'} = \underbrace{\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}}_{b'}.$$

Well, the "solution" to this system of three equations and three unknows is the point where the three columns intersect (in this case a point).

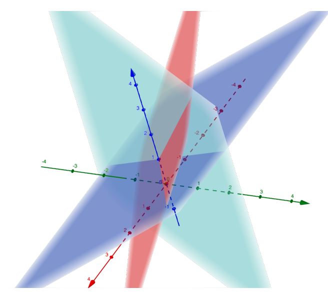


Figure 2. Row picture of $A' \cdot X' = b'$. Notice how with the added dimension, things get a little bit more complicated to draw and visualize with the row picture.

Subsection 1.3

Column Picture

1.3.1 Two Dimensions

Looking at the system of lienar equations:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned}$$

but this time, we look at the columns, giving us the equation in matrix form:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

This is more than a simple change of notation; written this way, we see that we want to scale by x vector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and scale by y vector $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ such that this combined scaling and adding of vectors gives us $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$.

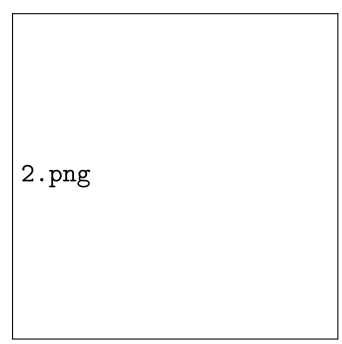


Figure 3. Column picture. We see that the *linear combination* of column 1 and column 2, such that we get vector $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ is one of column 1 plus two of column 2.

Definition 3

Linear Combination
Simply stated, a linear combination is just the adding and scaling of vectors (as seen above). More formally:

If $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k$ is a collection of vectors in $\mathbb{R}^n$, then a linear combination of the $\vec{\mathbf{v}}_i$ is a vector $\vec{\mathbf{w}}$ of the form

$$\vec{\mathbf{w}} = \sum_{i=1}^{k} a_i \vec{\mathbf{v}}_i \quad \text{for any scalars } a_i.$$

Now that we see that in this particular case, we are able to get to $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ by taking one of column 1 and two of column 2, but using just these two vectors, what is the set of vectors that we are able to create? Well, in this case we are able to fill the whole plane.

1.3.2 Three Dimensions

Given the system of three equations and three unknowns:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y - z &= -1 \\ -3y + 4z &= 4 \end{aligned}$$

then we can write this in column form as:

$$x \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} + z \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}.$$

For clarity let

$$\vec{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}, \text{ and } b' = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}.$$

We can clearly see that the linear combination such that $a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 = b'$ is

$$0\vec{v}_1 + 0\vec{v}_2 + 1\vec{v}_3 = b'.$$

The question is though, what is the set of points we are able to reach with various linear combinations $a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3$? Well, in this case the answer is all points in $\mathbb{R}^3$.

Subsection 1.4

Linear Independence and Span

More generally, can you solve $Ax = b$ for every b? That is do the linear combinations of the columns fill $3D$ space? Well, in what situations wouldn't this happen. Well given three columns, what are the different sets of b we can reach:

- Well if A is the zero matrix, then the set of b's would be $\vec{0}$.
- If all columns in A are vectors that lie on the same line, then their combinations would all lie on that line and so we could only reach points that lie on that line.
- If all columns were vectors lie on the same plane (but not on a single line) then there combinations would lie on that plane and so we could reach any point on that plane.
- If all columns are are vectors which each add a dimension, then all points in $\mathbb{R}^3$ can be reached.

More generally, if we have n columns, then all points in $\mathbb{R}^n$ can be reached only if each column (vector) is not a linear combination of the other columns (and it isnt the zero vector). That is, each column must be providing fundementally new information.

Definition 4

Linear Independence
A set of vectors $\vec{v}_1, \ldots, \vec{v}_k$ is linearly independent if the only solution to

$$a_1 \vec{\mathbf{v}}_1 + a_2 \vec{\mathbf{v}}_2 + \cdots + a_k \vec{\mathbf{v}}_k = \vec{\mathbf{0}} \quad \text{is} \quad a_1 = a_2 = \cdots = a_k = 0.$$

An equivalent definition is that vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k \in \mathbb{R}^n$ are linearly independent if and only if a vector $\vec{\mathbf{w}} \in \mathbb{R}^n$ can be written as a linear combination of those vectors in at most one way:

$$\sum_{i=1}^{k} x_i \vec{\mathbf{v}}_i = \sum_{i=1}^{k} y_i \vec{\mathbf{v}}_i \quad \text{implies} \quad x_1 = y_1, x_2 = y_2, \ldots, x_k = y_k.$$

Yet another equivalent definition is to say that $\vec{\mathbf{v}}_i$ are linearly independent if none of the $\vec{\mathbf{v}}_i$ is a linear combination of the others.

Definition 5

Span
Span is simply the set of points that can be reached through different linear combinations of the columns of A. More formally:
Given $S = \{\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k\}$ then the *span* of S is the set of linear combinations $a_1\vec{\mathbf{v}}_1 + \cdots + a_k\vec{\mathbf{v}}_k$. It is denoted $\operatorname{Span}(\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k)$.

Subsection 1.5

Matrix Vector Multiplication

First, what exactly is a matrix? Well, graphically, we can think of it as encoding a linear transformation. For example, given a general 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then each column just describes the new positions of the basis vectors. So then what does it mean to multiply a matrix by a vector? Well, recall that a vector depends upon the basis vectors. That is, the vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is really just a linear combination:

$$\vec{v} = v_1\hat{\imath} + v_2\hat{\jmath}.$$

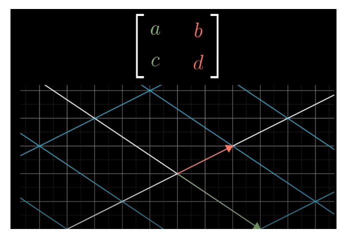


Figure 4. This general 2×2 matrix just describes the new positions of the basis vectors $\hat{\imath}$ and $\hat{\jmath}$.

With this knowledge, then matrix vector multiplication really just defines where the vector will be after the change of coordinates defined by the matrix. We can think of matrix vector multiplication as the as just a linear combination with new basis vectors. Thus

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \begin{bmatrix} a \\ c \end{bmatrix} + v_2 \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} av_1 + b_v 2 \\ cv_1 + dv_2 \end{bmatrix}.$$

That is, matrix vector multiplication just describes how the linear transformation $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ affects vector $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

Section 2

Lecture 2: Elimination with Matrices

As we mentioned in the last lecture, one of the fundemental uses of linear algebra is in solving systems of linear equations; given some equation

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

we want to find the the vector $\vec{x}$ such that $A\vec{x} = \vec{b}$. One of the ways we do this is via row reduction.

Subsection 2.1

Row Reduction

Let's say we want to solve the following system of linear equations:

$$\begin{aligned} 2x + y + 3z &= 1 \\ x - y &= 1 \\ 2x + z &= 1. \end{aligned}$$

Well, we could solve this using algebra; we could set $x = 1 + y$ and then substitute for x. Then we could set z in terms of y and substitute, and solve for z, then go back and substitute the known value for z into and solve that way. Instead, we can make this process more systematic by placing everything in matrices and performing matrix multiplication:

$$\begin{array}{cc} & \overbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}^{\vec{x}} \\ \underbrace{\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}}_{A} & \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\vec{\mathbf{b}}} \end{array}$$

For simplicity, we tend to omit $\vec{x}$ and combine A and $\vec{b}$ into a single matrix:

$$\begin{bmatrix} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{bmatrix}.$$

What exactly is row reduction? Well, the idea is we can use row operations to have one row in terms of z, another in terms of y, z and another in terms of x, y, z, and then back substitute to solve the equation.

Definition 6 **Row Operations**
A *row operation on a matrix is one of the three operations*:

1. Multiplying a row by a nonzero number
2. Adding a multiple of a row onto another row

3. Exchanging two rows

2.1.1 Example 1

Let's say we want to solve the following system of linear equations:

$$\begin{aligned} x + 2y + z &= 2 \\ 3x + 8y + z &= 12 \\ 4y + z &= 2. \end{aligned}$$

The first thing we will do is put it in a matrix:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

We will then take the following steps:

1. Find the first column that is not all 0's; call this the first pivotal column and call its first nonzero entry a pivot. If the pivot is not in the first row, move the row containing it to the first row position.

$$\begin{bmatrix} {\color{blue}1} & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

   We see that the first column is a pivotal column, and that the first nonzero entry is in the first row. Note the pivot in blue.

2. Divide the first row by the pivot, so that the first pivotal column is 1.
   We already have a pivotal 1, so no work to be done.

3. Add appropriate multiples of the first row to the other rows to make all other entries of the first pivotal column 0. Performing $R_2 - 3R_1$ we get:

$$\begin{bmatrix} {\color{blue}1} & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

4. Choose the next column that contains at least one nonzero entry beneath the first row, and put the row containing the new pivot in second row position. Make the pivot a pivotal 1: divide by the pivot, and add appropriate multiples of this row to the other **lower** rows, to make all other entries below the pivotal 1 of this column 0.
   We have a pivotal 2 in the second row:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & {\color{blue}2} & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

   Dividing R_2 by 2 to make it a pivotal 1, we get:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & {\color{blue}1} & -1 & 3 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

Then performing $R_3 - 4R_2$ we get:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 5 & -10 \end{bmatrix}.$$

5. Repeating the process, we divide R_3 by 5 to get a pivotal 1 in R_3:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}.$$

We thus have

$$z = -2, \quad y - z = 3, \quad x + 2y + z = 2.$$

Back substituting we get:

$$x = 2, y = 1, z = -2.$$

SUBSECTION 2.2
Matrix Multiplication

How do we multiply two matrices? Well first we should ask, is it possible? If A and B are two matrices, then it is only possible to perform AB if B has the same number of rows as A has columns. To multiply AB we often first write the matrix A then above and to the right matrix B, and in the space below we place the values of our new matrix $[AB]$.

$$\begin{array}{cc} & \begin{bmatrix} \quad B \quad \end{bmatrix} \\ \begin{bmatrix} A \end{bmatrix} & \begin{bmatrix} \quad AB \quad \end{bmatrix} \end{array}$$

Figure 5. When multiplying matrices A and B we often write them down in this manner. This makes it much more convenient to compute the values of the matrix $[AB]$.

2.2.1 Example 2

Let's look at how the following matrix multiplication was performed:

$$\begin{array}{cc} & \overbrace{\begin{bmatrix} 1 & 4 & -2 \\ 3 & 0 & 2 \end{bmatrix}}^{B} \\ \underbrace{\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}}_{A} & \underbrace{\begin{bmatrix} -1 & 8 & -6 \\ 9 & 12 & -2 \end{bmatrix}}_{AB} \end{array}$$

To compute the value of cell $(1,1)$, first row and first column we take the dot product of the first row of A and the first column of B:

$$[2\ -1] \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = (2 \cdot 1) + (-1 \cdot 3) = 2 - 3 = -1.$$

To compute the value of cell $(1,2)$, first row and second colum, we take the dot product of the first row of A with the second column of B:

$$[2\ -1] \cdot \begin{bmatrix} 4 \\ 0 \end{bmatrix} = (2 \cdot 4) + (-1 \cdot 0) = 8.$$

You then repeat the process for all the subsequent cells of AB. You can thus see why, to be able to perform matrix multiplication you B has to have the same number of rows as A has columns; if there was a mismatch, computing the dot product would be impossible.

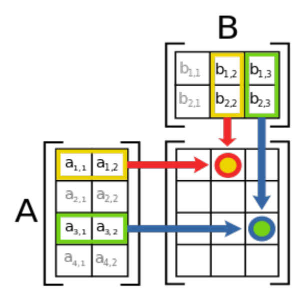


Figure 6. A visual representation of how to compute the value of each cell.

Definition 7

Matrix Multiplication
If A is an $m \times n$ matrix whose (i,j) th entry is $a_{i,j}$, and B is an $n \times p$ matrix whose (i,j) th entry is $b_{i,j}$, then $C = AB$ is the $m \times p$ matrix with entries

$$\begin{aligned} c_{i,j} &= \sum_{k=1}^{n} a_{i,k} b_{k,j} \\ &= a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \cdots + a_{i,n} b_{n,j}. \end{aligned}$$

2.2.2 Matrix-Column Multiplication

Just like matrix-vector multiplication is just a linear combination of columns, this idea exapnds to matrix-matrix multiplication. So for example if you wanted to multiply

$$\underbrace{\begin{bmatrix} \text{C} & \text{C} & \text{C} \\ \text{o} & \text{o} & \text{o} \\ \text{l} & \text{l} & \text{l} \\ \text{u} & \text{u} & \text{u} \\ \text{m} & \text{m} & \text{m} \\ \text{n} & \text{n} & \text{n} \\ & & \\ 1 & 2 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} 1C_1 + 2C_2 + 3C_3 & 4C_1 + 5C_2 + 6C_3 \end{bmatrix}}_{B}.$$

then you would have an $n \times 2$ matrix (where n is the number of rows of matrix A)

2.2.3 Row-Matrix Multiplication

We have already seen how matrix-vector multiplication is just a linear combination of columns (see 1.5), well row-matrix multiplication is just a linear combination of rows. So for example if you multiply

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

the result is the following linear combination of rows:

$$\begin{aligned} \begin{bmatrix} 3 & -2 & 0 \end{bmatrix} \begin{bmatrix} 3 & -2 & 0 \\ 2 & 1 & 2 \\ 0 & 4 & 3 \end{bmatrix} &= 3\begin{bmatrix} 3 & -2 & 0 \end{bmatrix} + 2\begin{bmatrix} 2 & 1 & 2 \end{bmatrix} + 0\begin{bmatrix} 0 & 4 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 9 & -6 & 0 \end{bmatrix} + \begin{bmatrix} 4 & 2 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 13 & -4 & 4 \end{bmatrix}. \end{aligned}$$

This idea expands when doing matrix-matrix multiplication. For example let's say you wanted to multiply

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \text{Row 1} \\ \text{Row 2} \\ \text{Row 3} \end{bmatrix}}_{A}$$

then the result would be

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \text{Row 1} \\ \text{Row 2} \\ \text{Row 3} \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1(R_1) + 2(R_2) + 3(R_3) \\ 4(R_1) + 5(R_2) + 6(R_3) \end{bmatrix}}_{B}.$$

2.2.4 Identity Matrix

Multiplying matrix A by the indentity matrix I returns back the matrix A. So it is just like multiplying a number by 1.

Definition 8 **Identity matrix**
The identity matrix I_n is the $n \times n$-matrix with 1's along the main diagonal (the diagonal from top left to bottom right) and 0's elsewhere.

Thinking in terms of row matrix multiplication this makes sense. If A is an $n \times m$ matrix then

$$AI_n = A = I_m A.$$

2.2.5 Scaling Columns of matrix

We can also modify the identity matrix to return a matrix that has a column or row scaled by some constant. For example let's say that we started out with a matrix A and want to multiply by some matrix $D_1(2)$ (Dilation Matrix) such that we get some matrix B whose first column is double the first column of A. Well than, we can use the matrix column definition of matrix multiplication to get the result we want; all we have to do is set the first element of D, $D_{1,1} = 2$, and the rest of the main diagonal equal to one, with everything else equal to zero:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{D_1(2)} = \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}}_{B}.$$

2.2.6 Switching Columns of Matrix

Can we use matrix-matrix multiplication to switch the columns of a matrix? Yes. Since we are dealing with columns we will use the matrix-column multiplication definition of matrix matrix multiplication. Let's say we want to multiply matrix A by some matrix X such that columns 1 and 4 are switched:

$$\underbrace{\begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} 7 & 3 & 5 & 1 \\ 8 & 4 & 6 & 2 \end{bmatrix}}_{B}.$$

Note that since we are using the matrix-column definition we just have to turn on the column that we want in matrix B.

2.2.7 Scaling Rows of Matrix

For this, instead of using the matrix-column definition, we will use the row-matrix definition. If we want to scale the third row of matrix A by 3 then we do the following:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{X} \cdot \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 21 & 24 & 27 \\ 10 & 11 & 12 \end{bmatrix}}_{B}.$$

Note that in matrix X we just have to turn on the rown that we want.

2.2.8 Properties of Matrix Multiplication

- **Not commutative**. Not only is matrix multiplication not commutative, if you can perform AB it might not even be possible to perform BA. For example if A is a 3×3 matrix and B is a 2×3 matrix, you can perform AB but not BA.

- **Associative**. $A(BC) = (AB)C$.
- **Distributative over addition**. $A(B+C) = AB+AC$ and $(B+C)A = BA+CA$.

2.2.9 Shear Matrix

INPUT CONTENT HERE.

2.2.10 Triangle Matrix

Definition 9

Triangle Matrix
A square matrix is called lower triangular if all the entries above the main diagonal are zero. Similarly, a square matrix is called upper triangular if all the entries below the main diagonal are zero.

$$L = \begin{bmatrix} \ell_{1,1} & & & & 0 \\ \ell_{2,1} & \ell_{2,2} & & & \\ \ell_{3,1} & \ell_{3,2} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n,1} & \ell_{n,2} & \dots & \ell_{n,n-1} & \ell_{n,n} \end{bmatrix}$$

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \dots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \dots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$

Figure 7. Lower triangular matrix L and upper triangular matrix U.

SUBSECTION 2.3

Elimination Matrix

Recall example 1 2.1.1. In that we performed various row operations to transform our matrix and eliminate different cells. Can we describe those elimination steps as matrices? For example, what is the matrix E_{21} that elinates cell $2,1$ (row 2, column 1), that brings about the follwing transformation

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}.$$

Note that row 1 and row 3 are left unchanged and only row 2 (in blue) is changed. That is we want to find the matrix E_{21} such that

$$\begin{matrix} & \begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix} \\ \underbrace{\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}}_{E_{21}} & \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix} \end{matrix}.$$

Here is where thinking in terms of row-matrix multiplication is really helpful (see 2.2.3); we want the first row of our answer to be $1R_1 + 0R_2 + 0R_3$ (that is, left unchanged). Similarily row 3 is left unchanged so we have E_{21}:

$$\begin{bmatrix} 1 & 0 & 0 \\ ? & ? & ? \\ 0 & 0 & 1 \end{bmatrix}.$$

To get the middle row (in blue) we perform the following linear combination of rows: $-3R_1 + 1R_2 + 0R_3$, and so we can now fill out the rest of the cells of E_{21}:

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

What does the subscript of E_{21} signify? Well, matrix E_{21} is the matrix that eliminates row 2 column 1 after performing matrix-matrix multiplication. Thinking about example

1 more broad picture, what are the steps (in terms of multiplication of Elimination matrices) that converts starting matrix A into an upper triangular matrix U?

$$\underbrace{\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}}_{A} \longrightarrow \underbrace{\begin{bmatrix}1&2&1\\0&2&-2\\0&0&5\end{bmatrix}}_{U}.$$

First note a couple changes, we have not included $\vec{b}$ in our matrix and we have not changed the pivots to be pivotal 1's. Well, what are the cells that we want to eliminate? We want to eliminate $(2,1)$ and $(3,2)$. Thus we have $E_{32}(E_{21}A) = U$:

$$\underbrace{\begin{bmatrix}1&0&0\\0&1&0\\0&-2&1\end{bmatrix}}_{E_{32}}\left(\underbrace{\begin{bmatrix}1&0&0\\-3&1&0\\0&0&1\end{bmatrix}}_{E_{21}}\underbrace{\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}}_{A}\right) = \underbrace{\begin{bmatrix}1&2&1\\0&2&-2\\0&0&5\end{bmatrix}}_{U}.$$

Due to the associativity of matrix multiplication we can move the parentheses, multiplying the elimination matrices first; in this way, we can encode all the steps necessary to transform go from $A \longrightarrow U$. Thus we have:

$$\underbrace{\begin{bmatrix}1&0&0\\-3&1&0\\6&-2&1\end{bmatrix}}_{E}\underbrace{\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}}_{A} = \underbrace{\begin{bmatrix}1&2&1\\0&2&-2\\0&0&5\end{bmatrix}}_{U}.$$

What if we wanted to change the pivotal $1, 2$ and 5 to be all pivotal ones? Well, we can scale the rows using matrices as well. Doing this we have

$$\underbrace{\begin{bmatrix}1&0&0\\-\frac{3}{2}&\frac{1}{2}&0\\\frac{6}{5}&-\frac{2}{5}&\frac{1}{5}\end{bmatrix}}_{E'}\underbrace{\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}}_{A} = \underbrace{\begin{bmatrix}1&2&1\\0&1&-1\\0&0&1\end{bmatrix}}_{U'}.$$

Subsection 2.4

Permutation Matrix

We have shown that we can use matrix multiplication to perform 2 of the 3 row operations (see 6), but what if we want to exchange two rows, can we encode that in a matrix? Well yes, but what would that look like? Let's say we want to perform matrix multiplication such that

$$\begin{bmatrix}a&b\\c&d\end{bmatrix} \longrightarrow \begin{bmatrix}c&d\\a&b\end{bmatrix}.$$

Well thinking in terms of row-matrix multiplication we want the top row to be $0R_1+1R_2$ and the bottom row to be $1R_1 + 0R_2$. Therefore the following matrix multiplication would perform the task:

$$\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix} = \begin{bmatrix}c&d\\a&b\end{bmatrix}.$$

What if instead we wanted to exchange columns? So for example we want to go from

$$\begin{bmatrix}a&b\\c&d\end{bmatrix} \longrightarrow \begin{bmatrix}b&a\\d&c\end{bmatrix}.$$

Well in this case, we don't want to think in terms of row-matrix multiplication but instead matrix-column multiplication. To do this, we put the multiplying matrix on the right. Thus, we want the first column to be $0C_1 + 1C_2$ and the second column to be $1C_1 + 0C_2$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}.$$

What if instead you wanted to go from

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} \longrightarrow \begin{bmatrix} c & d \\ a & b \\ e & f \end{bmatrix}?$$

Well, you would perform the following oeration:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \\ e & f \end{bmatrix}.$$

Notice that the relationship with the identity matrix; to switch rows 1 and 2 we exchanged rows 1 and 2 of the identity matrix.

Definition 10 **Permutation Matrix**
Let A be some $m \times n$ matrix. If we want to exchange rows i and j of matrix A using matrix multiplication:

$$P_{ij} A$$

then the *permuation matrix* P_{ij} is the identity matrix I_n but with rows i and j switched.

Subsection 2.5

Example 3 – Modelling Gravity

Let's say we want to use calculte the force of gravity while modelling the path a ball takes and the height of the ball at various times. Collecting the following information we have

Time	Height
3	3
3.3	2.559
3.6	1.236

Recall that the hieght of a falling ball at time t can be modelled by the following quadratic equation:

$$h(t) = a_0 \cdot 1 + a_2 t + a_3 t^2.$$

Given our data, we can thus have the following three equations:

$$\begin{aligned} h(3) &= a_0 + a_2(3) + a_3 9 = 3 \\ h(3.3) &= a_0 + a_2(3.3) + a_3(10.89) = 2.559 \\ h(3.6) &= a_0 + a_2(3.6) + a_3(12.96) = 1.236. \end{aligned}$$

Putting this in matrix form we have

$$\underbrace{\begin{bmatrix} 1 & 3.0 & 9.00 \\ 1 & 3.3 & 10.89 \\ 1 & 3.6 & 12.96 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 3.000 \\ 2.559 \\ 1.236 \end{bmatrix}}_{\vec{b}}.$$

We now need to form an upper triangle matrix. We already have a pivotal 1 in the first column, so using row-matrix multiplication we use an elimination matrix to zero out the rest of column 1:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}}_{E_{C_1}} \cdot \underbrace{\begin{bmatrix} 1 & 3.0 & 9.00 \\ 1 & 3.3 & 10.89 \\ 1 & 3.6 & 12.96 \end{bmatrix}}_{A} = \begin{bmatrix} 1 & 3.0 & 9.00 \\ 0 & 0.3 & 1.89 \\ 0 & 0.6 & 3.96 \end{bmatrix}.$$

Multiplying the second row by 3 to form a pivotal 1 and then performing $1R3 - 0.6R2$ using the row-matrix multiplication to zero out column 2 we have:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{10}{3} & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_{C_2}} \cdot \begin{bmatrix} 1 & 3.0 & 9.0 \\ 0 & 0.3 & 1.89 \\ 0 & 0.6 & 3.96 \end{bmatrix} = \begin{bmatrix} 1 & 3.0 & 9.0 \\ 0 & 1.0 & 6.3 \\ 0 & 0 & 0.18 \end{bmatrix}.$$

Note that in Matrix E_{C_2}, the reason we have -2 in row 3 column 2 is that -0.6 of $R2$. But realize that in this current matrix we are modifying $R2$ (multiplying by $\frac{10}{3}$ such that there is a pivotal 1). Thus we have to keep note of that we want -0.6 of $\frac{10}{3}$ of $R2$, which when combined gives us $-0.6 \cdot \frac{10}{3} = -2$. This is combining the matrix that creates the pivotal 1, and the matrix that eliminates row 3 column 2:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -0.6 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{10}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{10}{3} & 0 \\ 0 & -2 & 1 \end{bmatrix}.$$

Now all we have to do is turn 0.18 into 1, giving us the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{50}{9} \end{bmatrix} \cdot \begin{bmatrix} 1 & 3.0 & 9.0 \\ 0 & 1.0 & 6.3 \\ 0 & 0 & 0.18 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 3.0 & 9.0 \\ 0 & 1.0 & 6.3 \\ 0 & 0 & 1 \end{bmatrix}}_{A'}.$$

Combining all these steps into one elimination matrix we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{50}{9} \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{10}{3} & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_{C_2}} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}}_{E_{C_1}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -\frac{10}{3} & \frac{10}{3} & 0 \\ \frac{50}{9} & -\frac{100}{9} & \frac{50}{9} \end{bmatrix}}_{E}.$$

We also have to multiply $E \cdot \vec{B}$ to get the correct values now that we have changed A into an upper triangle matrix. This gives us

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -\frac{10}{3} & \frac{10}{3} & 0 \\ \frac{50}{9} & -\frac{100}{9} & \frac{50}{9} \end{bmatrix}}_{E} \cdot \underbrace{\begin{bmatrix} 3.000 \\ 2.559 \\ 1.236 \end{bmatrix}}_{\vec{b}} = \begin{bmatrix} 3 \\ -1.47 \\ -4.9 \end{bmatrix}.$$

We now have our full equation

$$\underbrace{\begin{bmatrix} 1 & 3.0 & 9.0 \\ 0 & 1.0 & 6.3 \\ 0 & 0 & 1 \end{bmatrix}}_{A'} \cdot \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 3 \\ -1.47 \\ -4.9 \end{bmatrix}}_{\vec{b}'}$$

Then by back substitution we have

$$a_2 = -4.9 \quad \text{and} \quad a_1 = 29.4 \quad \text{and} \quad a_0 = -41.1.$$

Note that you could also take the elimination matrix E a step further and do the back-substitution via matrice operations. By inspection we see that the following row operations on A' will turn A' into the identity matrix:

$$\begin{bmatrix} 1 & -3 & -9 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -6.3 \\ 0 & 0 & 1 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 1 & 3 & 9 \\ 0 & 1 & 6.3 \\ 0 & 0 & 1 \end{bmatrix}}_{A'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Well, since $E \cdot A = A'$. Then we can create a super elimination matrix E^* (because stars are always better) by just combining the matrices:

$$\begin{bmatrix} 1 & -3 & -9 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -6.3 \\ 0 & 0 & 1 \end{bmatrix} \cdot E = \underbrace{\begin{bmatrix} 66 & -120 & 55 \\ \frac{-115}{3} & \frac{220}{3} & -35 \\ \frac{50}{9} & \frac{-100}{9} & \frac{50}{9} \end{bmatrix}}_{E^*}.$$

Now, all we have to do is multiply E^* by $\vec{b}$ to get the solutions a_0, a_1, a_2:

$$\underbrace{\begin{bmatrix} 66 & -120 & 55 \\ \frac{-115}{3} & \frac{220}{3} & -35 \\ \frac{50}{9} & \frac{-100}{9} & \frac{50}{9} \end{bmatrix}}_{E^*} \cdot \underbrace{\begin{bmatrix} 3.000 \\ 2.559 \\ 1.236 \end{bmatrix}}_{\vec{b}} = \underbrace{\begin{bmatrix} -41.1 \\ 29.4 \\ -4.9 \end{bmatrix}}_{\vec{x}}.$$

The benefit of this is that assuming matrix A stays the same, you can change $\vec{b}$ as much as you want, and you just have to perform the one matrix-vector operation to get the solution. *Note: that E^* is actually the inverse of A.*

Definition 11 **Regular Matrix**
A square matrix is *regular* if it reduces to an upper-triangular matrix U with all nonzero pivots in diagonal entries using only shear matrices.

Section 3

Lecture 3: Matrice Inverses

What are inverse matrices and why do we care about them? Well, as we have hinted at above matrice inverses can be used to solve systems of linear equations. In addition they also give us a way to determine if vectors are linearly independent. First what is an inverse matrix and when is a matrix invertable?

Definition 12 **Invertible Matrix**
Let $A \in \mathbb{R}^{n \times n}$ be a square matrix. We say that A is *invertible* if and only if there exists a matrix $C \in \mathbb{R}^{n \times n}$ such that

$$A \cdot C = C \cdot A = I_n$$

where I_n is the $n \times n$ identify matrix. We call $C = A^{-1}$ the *inverse* of A. An invertible matrix is called non-singular while a matrix that is not invertible is called singular.

We mentioned that if matrice inverses can help us solve systems of linear equations. Why is this and how?

Theorem 1 **Solving Equations with Matrix Inverse**
If A has an inverse A^{-1}, then for any $\overrightarrow{\mathbf{b}}$ the equation $A\overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{b}}$ has a unique solution, namely $\overrightarrow{\mathbf{x}} = A^{-1}\overrightarrow{\mathbf{b}}$.

Proof: First verifying that $A^{-1}\vec{b}$ is a solution:

$$\begin{aligned}
A \cdot \vec{x} &= A \cdot (A^{-1} \cdot \vec{b}) && \textit{Let } \vec{x} = A^{-1}\vec{b} \\
&= (A \cdot A^{-1}) \cdot \vec{b} && \textit{Associativity of matrix multiplication} \\
&= I\vec{b} && \textit{Definition of inverse matrix} \\
&= \vec{b}. && \textit{Definition of identity matrix}
\end{aligned}$$

Now proving uniqueness:

$$\begin{aligned}
A\vec{x} &= \vec{b} \\
A^{-1}A\vec{x} &= A^{-1}\vec{b} && \textit{Left matrix multiplication} \\
I\vec{x} &= A^{-1}\vec{b} && \textit{Definition of inverse matrix} \\
\vec{x} &= A^{-1}\vec{b}. \quad \square
\end{aligned}$$

Source: Hubbard
We thus see that if A is invertible and we find its inverse we have found a way to compute the solution. In this way, by just finding the inverse matrix we find a way to compute solution even if $\vec{b}$ changes.

We have also mentioned that we can use matrix inverses to determine if vectors are linearly independent or not. Why is this?

Theorem 2 **Square matrices and linearly independent columns**
Given a square matrix $A \in \mathbb{R}^{n \times n}$ then A is invertible if and only if its columns are linearly independent.

Proof: We prove both directions of implication.

($\Rightarrow$) Fist we show that if A is invertible, then its columns are linearly independent.

First recall the definition of linear indipendence, see 4. Well, in terms of matrix-vector multiplication this means that the columns of A are linearly independent is equivalent to stating that $A\vec{x} = \vec{0}$ only when $\vec{x} = \vec{0}$. Thus to show linear independence we just have to prove that

$$A \text{ is invertible } \Rightarrow (A\vec{x} = \vec{0} \text{ only when } \vec{x} = \vec{0}).$$

Left multiplying both sides by A^{-1} (which we can do because A is invertible) we have

$$\begin{aligned} A\vec{x} &= \vec{0} \\ A^{-1}A\vec{x} &= A^{-1} \cdot \vec{0} \\ I\vec{x} &= \vec{0} \\ \vec{x} &= \vec{0}. \end{aligned}$$

This is what we needed to show.

($\Leftarrow$) Next, we prove that if the columns of A are linearly independent then A is invertible. Because the columns of A are linearly independent a solution exists for linear-systems equation $A\vec{x} = \vec{b}$. We haven't yet covered this fact, but we can prove this from the fact that n indipendent vectors span an n-dimensional vector space. For example, basis vectors, by necessity are linearly independent. We can thus rewrite our implication as

$$A\vec{x} = \vec{b} \text{ has a solution } \vec{x} \to A^{-1} \text{ exists.}$$

We want to show that there exists a matrix M such that $AM = MA = I$. Well because we know that $A\vec{x} = \vec{b}$ has a solution $\vec{x}$ for all $\vec{b}$, can we rewrite $AM = MA = I$ in terms of matrix-vector multiplication? Using matrix-column multiplication we rewrite $AM = I$ as

$$\begin{gathered} AM = I \\ \mathbf{A} \begin{bmatrix} | & | & & | \\ \vec{m}_1 & \vec{m}_2 & \cdots & \vec{m}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \vec{e_1} & \vec{e_2} & \cdots & \vec{e_n} \\ | & | & & | \end{bmatrix} \\ \begin{bmatrix} | & | & & | \\ \mathbf{A}\vec{m}_1 & \mathbf{A}\vec{m}_2 & \cdots & \mathbf{A}\vec{m}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \vec{e_1} & \vec{e_2} & \cdots & \vec{e_n} \\ | & | & & | \end{bmatrix} \end{gathered}$$

where $\vec{m}_i$ and $\vec{e}_i$ are the columns of M and the identity matrix respectively. Well we know that there exist vectors $\vec{m}_1, \ldots, \vec{m}_n$ such that

$$A\vec{m}_i = \vec{e}_i \forall i \in 1, 2, \ldots, n$$

because we know that $A\vec{x} = \vec{b}$ always has a solution. $\square$
Source: Berekley Notes

Subsection 3.1
Elementary Matrices

In our example on gravity, we performed many row operations to create our elimination matrices. The idea behind elementary matrices is that any row operation on a matrix A by multiplying A on the left by an elementary matrix. There are three types of elementary matrices, each corresponding to a row operation; and each defined in terms of the identity matrix (diagonal all ones, everything else zero). See 2.2.5 and 6 for more information on row and column operations. Source: Hubbard

3.1.1 Row-Multiplication

This is the elementary matrix that scales row i by x, and can be expressed as $E_1(i, x)$, E_1 because it is the first elementary matrix. For example, if we have matrix

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

then the elementary matrix to multiply row 2 by 5 would be

$$\begin{aligned} E_1(2,5) \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 15 & 20 \end{bmatrix}. \end{aligned}$$

Specifically, the row-multiplication elementary matrix $E_1(i, x)$ is the square matrix where all nondiagonal entries are 0, and every entry on the diagonal is 1 except for the (i, i)th entry, which is $x \neq 0$. Source: Hubbard
This matrix can be formed as

$$E_1(i, x) = I_n + x e_i e_i^T$$

where I_n is the $n \times n$ square matrix and e_i is the ith column of the identiy matrix. Source: Jeff's Lectures Also, note that the inverse of this elementary matrix is

$$(E_1(i, x))^{-1} = E_1(i, \frac{1}{x}).$$

3.1.2 Row Addition

This is the elementary matrix that replaces a row of matrix A with the sum of that row and a multiple of another. This is represented as $E_2(i, j, x)$ where i is the row we will be replacing, j is the row we will be taking a multiple of, and x is the multiple of the row. For example, if we wanted to zero the 3 in matrix A above we would subtract 3 of row 1 from row 2:

$$\begin{aligned} E_2(2,1,-3) \cdot A &= \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}. \end{aligned}$$

Specifically the row addition elementary matrix $E_2(i, j, x)$, for $i \neq j$, is the square matrix with all diagonal entries 1, and all other entries 0 except for the (i, j)th, which is x. Source: Hubbard
This matrix can be formed as

$$E_2(i, j, x) = I_n + x e_i \vec{e}_j^T$$

where $\vec{e}_i$ is the ith column of the identity matrix and e_j^T is the transpose of the jth column of the identity matrix. Note that the inverse of this elementary matrix is

$$E_2(i, j, x)^{-E} = 1_2(i, j, -x).$$

3.1.3 Row Switching

This is the elementary matrix $E_3(i,j)$, $i \neq j$ that switches row i and row j of some matrix A. So for example if we wanted to switch rows 1 and 2 from the A above, then we would have

$$\begin{aligned} E_3(i,j) \cdot A &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

Specifically the row switching matrix $E_3(i,j)$, $i \neq j$, is the square matrix where the entries i,j and j,i are 1, as are all entries on the diagonal except i,i and j,j. Source: Hubbard We can more formally define the row switching matrix as

$$P_{i,j} = \vec{e}_i \vec{e}_j^T + \vec{e}_j \mathbf{e}_i^T + \sum_{\substack{k=1 \\ k \neq i,j}}^{n} \vec{e}_j \vec{e}_j^T$$

where $n \in \mathbb{N}$ and $\vec{e}_i$ is the ith column of the identity matrix and $\vec{e}_j$ is the jth column of the identity matrix.

Subsection 3.2

Elementary Matrices are Invertible

3.2.1 Shear Matrices

Given the shear matrix $S_{ik}(c)$, then its inverse is $(S_{ik}(c))^{-1} = S_{ik}(-c)$. How can we prove this? Well recall that a shear matrix can be constructed in the following manner:

$$S_{ik}(c) = I_n + ce_i e_k^T.$$

Well then we simply just multiply the shear by its (supposed) inverse, and we will see that the result is the identity matrix. To see this note that

$$\begin{aligned} S_{ik}(c) \cdot S_{ik}(-c) &= \left(I_n + c\vec{e}_i \vec{e}_k^T\right) \cdot \left(I_n - c\vec{e}_i \vec{e}_k^T\right) \\ &= I_n \cdot \left(I_n - c\vec{e}_i \vec{e}_k^T\right) + c\vec{e}_i \vec{e}_k^T \cdot \left(I_n - c\vec{e}_i \vec{e}_k^T\right) \\ &= I_n - c\vec{e}_i \vec{e}_k^T + c\vec{e}_i \vec{e}_k^T \cdot I_n - c^2 \left(\vec{e}_i \vec{e}_k^T\right) \cdot \left(\vec{e}_i \vec{e}_k^T\right) \\ &= I_n - c\vec{e}_i \vec{e}_k^T + c\vec{e}_i \vec{e}_k^T - c^2 \vec{e}_i \left(\vec{e}_k^T \vec{e}_i\right) \vec{e}_k^T \\ &= I_n - c^2 \underbrace{e_i}_{n\times 1} \cdot (\underbrace{e_k^T}_{1\times n} \underbrace{e_i}_{n\times 1}) \underbrace{e_k^T}_{1\times n} \\ &= I_n - c^2 \underbrace{e_i}_{n\times 1} \underbrace{[0]}_{1\times 1} \underbrace{e_k^T}_{1\times n} \\ &= I_n - c^2 \underbrace{e_i}_{n\times 1} \underbrace{[0]}_{1\times n} \\ &= I_n - c^2 \underbrace{[0]}_{n\times n} \\ &= I_n - \underbrace{[0]}_{n\times n} \\ &= I_n. \end{aligned}$$

Question: I don't understand why $e_k^T e_i$ results in a scalar output.I would have thought that the result is the zero matrix in $\mathbb{R}^1$.

3.2.2 Permutation Matrices

How can we prove that $(P_{ik})^{-1} = P_{ik}^T$?

3.2.3 Dilation matrices

Insert proof here.

Subsection 3.3

The Invertible Matrix Theorem

Let $A \in \mathbb{R}^{n\times n}$. Then the following statements are equivalent:

1. There is a matrix $C \in \mathbb{R}^{n\times n}$ such that $CA = I_n$.
2. There is a matrix $D \in \mathbb{R}^{n\times n}$ such that $AD = I_n$.
3. A is an invertible matrix (A is nonsingular).
4. A is row equivalent to an upper triangular matrix with nonzero entries on the main diagonal.
5. A has n pivot positions.
6. The matrix equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = 0$.
7. The columns of A are linearly independent. In other words
$$\{A(:,1), A(:,2), \ldots, A(:,n)\}$$
is a linearly independent set of vectors.
8. The linear transformation $f(\mathbf{x}) = A\mathbf{x}$ is one-to-one.
9. The equation $A\mathbf{x} = \mathbf{b}$ has a unique solutions for all $\mathbf{b} \in \mathbb{R}^n$.
10. The columns of A span $\mathbb{R}^n$.
11. The linear transformation $f(\mathbf{x}) = A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
12. The matrix A^T is invertible.
13. The determinant of A is not zero: $\text{set}(A) \neq 0$.
14. The columns of A form a basis for $\mathbb{R}^n$
15. The column space of A is $\mathbb{R}^n : \text{Col}(A) = \mathbb{R}^n$.
16. The dimension of the column space of A is $n : \dim(\text{Col}(A)) = n$
17. The rank of A is $n : \text{rank}(A) = n$.
18. The null space of A is $\{\mathbf{0}\} : \text{Null}(A) = \{\mathbf{0}\}$.
19. The dimension of the null space of A is $0 : \dim(\text{Nulll}(A)) = 0$.
20. The orthogonal complement of the column space of A is $\{\mathbf{0}\}$. We can write this as
$$(\text{Col}(A))^\perp = \{\mathbf{0}\}$$

21. The orthogonal complement of the null space of A is $\mathbb{R}^n$: We write this

$$(\text{Null}(A))^{\perp} = \mathbb{R}^n$$

22. The row space of A is $\mathbb{R}^n$: $\text{Row}(A) = \mathbb{R}^n$.
23. The number 0 is not an eigenvalue of A.
24. The matrix A has n non-zero singular values.

Source: Jeff's Notes, Lesson 14

Section 4

Lecture 4: LU Factorization

The idea behind LU Factorization (or really any factorization for that matter) is to decouple the computationally expensive factorization phase from the actual solving phase. Let's say you are solving some problem $A\vec{x} = \vec{b}$ well then since you have already done the factorization, you can solve the for $\vec{x}$ relativly quickly even if $\vec{b}$ keeps changing (which it often does). Source: StackExchange

What is LU factorization and how does it help us solve problems of the form $A\vec{x} = \vec{b}$? Well, $A = LU$ factorization takes decomposes A into a product of a lower triangle matrix L and an upper triangle matrix U. We have already seen how to form U as a sequence of elementary matrix operations (row additions) on A (where no row exchanges are required). For example if A is a 3×3 matrix then we would perform

$$E_{32}E_{31}E_{21}A = U$$

Where E_{ij} are the elementary elimination matrices responsible for zeroing out the values at position (i, j).

How do we find L? Well, recall that the inverse of a row-addition elementary matrix is really easy to find so we just have

$$\begin{aligned} E_{32}E_{31}E_{21}A &= U \\ A &= E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}U \\ &= LU. \end{aligned}$$

Notice that we took the inverses in reverse order.

Now that we have $A = LU$, how can we use that information to solve a system of linear equations? Well by substitution we have

$$\begin{aligned} A\vec{x} &= \vec{b} \\ LU\vec{x} &= \vec{b} \\ L(U\vec{x}) &= \vec{b} \\ L\vec{y} &= \vec{b}. && \textit{Let } \vec{y} = U\vec{x}\textit{, } \vec{y} \textit{ is a vector} \\ \vec{y} &= L^{-1}\vec{b}. \end{aligned}$$

We can easily solve for $\vec{y}$ through forward substitution. Then now that we have found $\vec{y}$ we use the equation $\vec{y} = U\vec{x}$ and solve for $\vec{x}$.

Definition 13 **LU Factorization without Pivoting**
Let $A \in \mathbb{R}^{n\times n}$ be a given, square, invertible matrix with non-zeros on the main

diagonal elements. An LU factorization of A is given by

$$A = LU$$

where $U \in \mathbb{R}^{n\times n}$ is upper-triangular with nonzero diagonal elements. Also, L is unit lower-triangular with all of its diagonal entries equal to 1 .
Source: Jeff's Notes, Lecture 15

SUBSECTION 4.1
Examples

4.1.1 Two by Two Matrix

Let's say we have the following linear systems problem

$$\underbrace{\begin{bmatrix} 2 & 4 \\ 4 & 10 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}_{\vec{b}}.$$

First finding the upper triangle matrix we have

$$\begin{aligned} \underbrace{\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}}_{E_2(2,1,-2)} \underbrace{\begin{bmatrix} 2 & 4 \\ 4 & 10 \end{bmatrix}}_{A} &= \underbrace{\begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}}_{U} \\ \underbrace{\begin{bmatrix} 2 & 4 \\ 4 & 10 \end{bmatrix}}_{A} &= \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}}_{U}. \qquad L = (E_2(2,1,-2))^{-1} \end{aligned}$$

Let $\vec{y} = U\vec{x}$, and so solving for $\vec{y}$ we have

$$\begin{aligned} L\vec{y} &= \vec{b} \\ \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}}_{L} \vec{y} &= \underbrace{\begin{bmatrix} 4 \\ 5 \end{bmatrix}} \\ \vec{y} &= \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ -3 \end{bmatrix}. \end{aligned}$$

Now solving for $\vec{x}$ we have

$$\begin{aligned} \vec{y} &= U\vec{x} \\ \begin{bmatrix} 4 \\ -3 \end{bmatrix} &= \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

By back substitution we have $x_1 = 5$ and $x_2 = -1.5$.

4.1.2 Three by Three Matrix

Let

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 9 & 14 \\ 7 & 16 & 26 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{\vec{b}}$$

be our liner systems problem. First finding U through row-addition we have:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_2(2,1,-4)} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 9 & 14 \\ 7 & 16 & 26 \end{bmatrix}}_{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 7 & 16 & 26 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix}}_{E_2(3,1,-7)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 7 & 16 & 26 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 2 & 5 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_2(3,2,-2)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 2 & 5 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{U}.$$

In short we have

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_2(3,2,-2)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix}}_{E_2(3,1,-7)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_2(2,1,-4)} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 9 & 14 \\ 7 & 16 & 26 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{U},$$

and thus taking the inverses of our elementary matrices to find L we have

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 9 & 12 \\ 7 & 16 & 26 \end{bmatrix}}_{A} = \underbrace{\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{(E_2(2,4,-4)^{-1})} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & 0 & 1 \end{bmatrix}}_{(E_2(3,1,-7))^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}}_{(E_2,3,2,-2))^{-1}}}_{L} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{U}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{U}.$$

Let $\vec{y} = U\vec{x}$ (a vector), and solving for $\vec{y}$ we have

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\vec{y}} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{\vec{b}}$$

which we can solve using forward substitution giving us

$$\begin{aligned} y_1 &= 1 \\ y_2 &= -2 \\ y_3 &= 0. \end{aligned}$$

Now using $\vec{y}$ to solve for $\vec{x}$ we have

$$\begin{aligned} \vec{y} &= U\vec{x} \\ \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} &= \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{U} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned}$$

Solving this via back substitution we have

$$\begin{aligned} x_1 &= 5 \\ x_2 &= -2 \\ x_3 &= 0. \end{aligned}$$

Subsection 4.2
Basic Facts

4.2.1 Inverse of a Matrix Product

Note that the inverse of a matrix product is

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof: Recall the definition of matrix inverse, see 12. We need to show that

$$(AB)(B^{-1}A^{-1} = I) \quad \text{and} \quad (B^{-1}A^{-1})(AB) = I.$$

First checking the multiplication on the right side:

$$\begin{aligned} (AB)(B^{-1}A^{-1} &= A(BB^{-1})A^{-1} \\ &= A(IA^{-1}) \\ &= AA^{-1} \\ &= I. \end{aligned}$$

Now checking the multiplication on the left side:

$$\begin{aligned} (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}(IB) \\ &= B^{-1}B \\ &= I. \quad \square \end{aligned}$$

4.2.2 Transpose of a Matrix Product

The transpose of a matrix product is

$$(AB)^T = B^T A^T.$$

Also note that

$$(A^{-1})^T = (A^T)^{-1}$$

SUBSECTION 4.3

Why $\boldsymbol{A = LU}$ and not $\boldsymbol{EA = L}$?

Recall our 3×3 linear systems problem from above:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_2(3,2,-2)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix}}_{E_2(3,1,-7)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_2(2,1,-4)} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 9 & 14 \\ 7 & 16 & 26 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{U}.$$

What happens when we form E from the product of our elimination matrices $E_2(3,2,-2) \cdot E_2(3,1,-7) \cdot E_2(2,1,-4)$? Calculating the product we have

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_2(3,2,-2)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix}}_{E_2(3,1,-7)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_2(2,1,-4)} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_2(3,2,-2)} \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix}$$

At this point, everything is going along smoothly; the -4 and -7 stayed in the same place, and so we can store the two elimination matrices in one matrix. But performing the next multiplication we have

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{E_2(3,2,-2)} \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}.$$

Now we have a 1 instead of -7, and so we can no longer store the three elimination matrices in one matrix, and thus the multiplication is more difficult to perform. What went wrong? When calculating the last row we had $-2R_2+R_3$ but the problem is that R_2 already has a value in the first column which messes with the -7 in the third row. But on the other hand, when calculating $L = (E_2(2,4,-4))^{-1} \cdot (E_2(3,1,-7))^{-1} \cdot (E_2(3,2,-1))^{-1}$ we have the nice property that when performing the multiplication we first fill out the bottom rows, so that the upper rows resemble the identity matrix and so nothing is messed with:

$$\begin{aligned} L &= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{(E_2(2,4,-4))^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & 0 & 1 \end{bmatrix}}_{(E_2(3,1,-7))^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}}_{E_2(3,2,-2)^{-1}} \\ &= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{(E_2(2,4,-4))^{-1}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix}. \end{aligned}$$

In this way we see that the lower triangle matrix L can easily be obtained from the elementary row addition matrices by just flipping the signs.

Subsection 4.4

Cost of $A = LU$ Elimination

Given an $n \times n$ matrix, what is the cost of factoring A into L and U? Recall, given the problem

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

we first want to find the matrix E such then when you perform $EA = U$, where E is the product of elementary matrices and U is an upper triangle matrix.

4.4.1 Cost of finding E

Let's say A is the 3×3 matrix

$$\begin{bmatrix} A_{00} & A_{01} & A_{02} \\ A_{10} & A_{11} & A_{12} \\ A_{20} & A_{21} & A_{22} \end{bmatrix}.$$

We first want to perform a series of multiplications and subtractions such that we $A_{10} = 0$ and $A_{20} = 0$. Well to do this, at each row we perform n multiplications and n subtractions. Because there are n rows this equates to

$$\begin{aligned} &n(n \text{ multiplications } + n \text{ subtractions}) \\ &n^2 \text{ multiplications } + n^2 \text{ subtractions}. \end{aligned}$$

We actually overestimated a little bit since we aren't changing the first row, so assuming we don't want pivotal 1's there will actually be $n^2 - n$ multiplications and $n^2 - n$ subtractions. But for the sake of simplicity, we'll go with our estimatimation of $2n^2$ for converting the first column to containing only one nonzero number.

After this step we have the following matrix

$$\begin{bmatrix} A_{00} & A_{01} & A_{02} \\ 0 & A'_{11} & A'_{12} \\ 0 & A'_{21} & A'_{22} \end{bmatrix}.$$

At this point we only really have to worry about the $(n-1) \times (n-1)$ matrix. For the same reasoning above, that will involve $(n-1)^2$ multiplications and $(n-1)^2$ subtractions. Thus, for any matrix we will have in total

$$\begin{aligned} (n^2 + n^2) + ((n-1)^2 + (n-1)^2) + \cdots + (1^2 + 1^2) &= 2n^2 + 2(n-1)^2 + \cdots + 2(1) \\ &= 2(n^2 + (n-1)^2 + \cdots + 1) \\ &= 2(\frac{1}{3}n(n + \frac{1}{2})(n+1)) \text{ operations.} \end{aligned}$$

When n gets to be large we can ignore the $\frac{1}{2}$ and 1, and thus we have about $\frac{2}{3}n^3$ operations.

Section 5

Lecture 8: Row Reduced Form

Up until now we have been solving problems of the form

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

where A is a square matrix. What about the more general case where $A \in \mathbb{R}^{m \times n}$?

Definition 14

The General Linear-Systems Problem
Let $m, n \in \mathbb{N}$. Let $A \in \mathbb{R}^{m \times n}$ be a given rectangular matrix and $\vec{b} \in \mathbb{R}^m$ be a given vector. Then the general linear-systems problem is to find an unknown vector $\vec{x} \in \mathbb{R}^n$ such that

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

As before, instead of solving the problem by finding the inverse of A directly, we want to first find the upper triangle matrix U so we can solve the easier problem of

$$U\vec{x} = \vec{y}.$$

This time, instead of simply being an upper triangle matrix, we transform A into an upper triangle matrix U in reduced row echelon form.

Definition 15

Row Echelon Form
Let $U \in \mathbb{R}^{m \times n}$ be a given matrix. We say that U is in row echelon form if and only if U satisfies the following two conditions

i. All zeros rows are below all nonzero rows

ii. The column index of the first nonzero entry in a row is larger than the column index of the first nonzero entry in any previous row.

iii. All entries in a column below an leading entry are zeros.

5.0.1 Matrices in Row Echelon Form

The following are matrices in row echelon form:

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 9 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 6 & 7 & 8 \end{bmatrix}.$$

5.0.2 Matrices not in Row Echelon Form

The following are matrices not in row echelon form:

$$\begin{bmatrix} 2 & 3 & 4 \\ 2 & 2 & 9 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 7 & 8 \end{bmatrix}.$$

Definition 16

Reduced Row Echelon Form
A matrix is in echelon form if and only if

1. In every row, the first nonzero entry is 1, called a pivotal 1.
2. The pivotal 1 of a lower row is always to the right of the pivotal 1 of a higher row.
3. In every column that contains a pivotal 1, all other entries are 0.
4. Any rows consisting entirely of 0 's are at the bottom.

Source: Hubbard

5.0.3 Matrices in Reduced Row Echelon Form

Here are some example matrices in reduced row echelon form. Note the underlying of the pivotal 1s:

$$\begin{bmatrix} \underline{1} & 0 & 0 & 3 \\ 0 & \underline{1} & 0 & -2 \\ 0 & 0 & \underline{1} & 1 \end{bmatrix}, \quad \begin{bmatrix} \underline{1} & 1 & 0 & 0 \\ 0 & 0 & \underline{1} & 0 \\ 0 & 0 & 0 & \underline{1} \end{bmatrix}, \quad \begin{bmatrix} 0 & \underline{1} & 3 & 0 & 0 & 3 & 0 & -4 \\ 0 & 0 & 0 & \underline{1} & -2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{1} & 2 \end{bmatrix}.$$

We see that being in row reduced echelon form does not mean that we are dealing with the identity matrix, but rather that the columns with a pivotal 1 do not have any other numbers in them. Note that unlike the identity matrix, not all columns must contain a pivotal 1.

5.0.4 Matrice not in Reduced Row Echelon Form

The following matrices are not in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 3 & 0 & -3 \\ 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Let's look at each matrix and say why not:

1. The pivotal 1 in row 3 is not to the right of the pivotal 1 in row 2. No put in echelon form, we just need to swap rows 2 and 3.
2. The first nonzero entry of row 2 is not a 1. We would just need to scale row 2 by $\frac{1}{2}$.
3. The row of all zeros is not on the bottom.
4. The first nonzero entry of row 2 is not 1, it is -1. Also, not row 3 has a pivotal 1, but that colum is not all zeros.

5.0.5 Gaussian Elimination

Here is the algorithm for putting a matrix $A \in \mathbb{R}^{m \times n}$ in reduced row echelon form:

1. Find the first column that is not all 0 's; call this the first pivotal column and call its first nonzero entry a pivot. If the pivot is not in the first row, move the row containing it to first row position.
2. Divide the first row by the pivot, so that the first entry of the first pivotal column is 1 .
3. Add appropriate multiples of the first row to the other rows to make all other entries of the first pivotal column 0 . The 1 in the first column is now a pivotal 1.
4. Choose the next column that contains at least one nonzero entry beneath the first row, and put the row containing the new pivot in second row position. Make the pivot a pivotal 1: divide by the pivot, and add appropriate multiples of this row to the other rows, to make all other entries of this column 0 .
5. Repeat until the matrix is in echelon form. Each time choose the first column that has a nonzero entry in a lower row than the lowest row containing a pivotal 1 , and put the row containing that entry directly below the lowest row containing a pivotal 1 .

Source: Hubbard

5.0.6 Row Reduction Example

Let's say we want to solve the following system of linear equations:

$$\begin{aligned} x + 2y + z &= 2 \\ 3x + 8y + z &= 12 \\ 4y + z &= 2. \end{aligned}$$

The first thing we will do is put the equation in matrix form:

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}.$$

We will then take the following steps:

1. Find the first column that is not all 0's; call this the first pivotal column and call its first nonzero entry a pivot. If the pivot is not in the first row, move the row containing it to the first row position.

$$\begin{bmatrix} \color{blue}{1} & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}.$$

   We see that the first column is a pivotal column, and that the first nonzero entry is in the first row. Note the pivot in blue.

2. Divide the first row by the pivot, so that the first pivotal column is 1.
   We already have a pivotal 1, so no work to be done.

3. Add appropriate multiples of the first row to the other rows to make all other entries of the first pivotal column 0. Performing $R_2 - 3R_1$ we get:

$$\begin{bmatrix} \color{blue}{1} & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}.$$

4. Choose the next column that contains at least one nonzero entry beneath the first row, and put the row containing the new pivot in second row position. Make the pivot a pivotal 1: divide by the pivot, and add appropriate multiples of this row to the other **lower** rows, to make all other entries below the pivotal 1 of this column 0.
   We have a pivotal 2 in the second row:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & \color{blue}{2} & -2 \\ 0 & 4 & 1 \end{bmatrix}.$$

   Dividing R_2 by 2 to make it a pivotal 1, we get:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & \color{blue}{1} & -1 \\ 0 & 4 & 1 \end{bmatrix}.$$

Then performing $R_3 - 4R_2$ we get:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 5 \end{bmatrix}.$$

5. Repeating the process, we divide R_3 by 5 to get a pivotal 1 in R_3:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

We thus have

$$z = -2, \quad y - z = 3, \quad x + 2y + z = 2.$$

Back substituting we get:

$$x = 2, y = 1, z = -2.$$

SUBSECTION 5.1

Solutions to the General Linear Systems Problem

The general solution to

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

takes the form

$$A[\vec{p} + \vec{z}] = \vec{b}$$

where $\vec{p}$ is a particular solution and $\vec{z}$ is the solution to $A\vec{x} = \vec{0}$ (which is a homogeneous equation). Why do we need $\vec{z}$ to describe the entire solution set? To make this more tangible let's look at a specific example.

5.1.1 Example 2

Let

$$\underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 8 \end{bmatrix}}_{\vec{b}}$$

be our linear systems problem. Clearly $\vec{x} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$ is a solution, but this is not the entire solution set. The entire solution set is described by

$$\vec{x} = \underbrace{\begin{bmatrix} 8 \\ 0 \end{bmatrix}}_{\vec{p}} + \underbrace{\begin{bmatrix} 0 \\ x_2 \end{bmatrix}}_{\vec{z}}$$

where x_2 is any real number. We thus see that there are infinitly many solutions, and thus we need $\vec{z}$ to account for this.

5.1.2 How Many Solutions?

Given a GLSP, there will either be no solutions, one solution, or infinitly many solutions, so which one is it?

Theorem 3 **Solutions to linear equations**
We can represent the linear systems problem $A\vec{x} = \vec{b}$ by the $m \times (n+1)$ matrix $[A|\vec{b}]$, which row reduces to $[U|\vec{y}]$. Then

1. If the row-reduced vector $\vec{y}$ contains a pivotal 1, the system has no solutions.
2. If $\vec{y}$ does not contain a pivotal 1, then solutions are uniquely determined by the values of the nonpivotal variables:
   a) If each column of U contains a pivotal 1 ($U = I$), the system has a unique solution.
   b) If at least one column of U is nonpivotal, there are infinitely many solutions: exactly one for each value of the nonpivotal variables.

Let's try and understand this better with some examples:

5.1.3 Example 3

We can represent the system of linear equations

$$\begin{aligned} 2x + y + 3z &= 1 \\ x - y + 0z &= 1 \\ x + y + 2z &= 1 \end{aligned}$$

as the matrix

$$\underbrace{\begin{bmatrix} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & 2 & 1 \end{bmatrix}}_{[A|\vec{b}]} \to \underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{[U|\vec{y}]}.$$

We thus see that the $\vec{y}$ contains a pivotal 1. Recall that for 1 to be pivotal, it means that it is the first nonzero entry in the row. Therefore, for $\vec{y}$ to contain a pivotal 1, the matrix is saying that $0 = 1$ is a solution (which of course is impossible).

5.1.4 Example 4

The system of linear equations

$$\begin{aligned} x + 2y + 5z &= 8 \\ x + y + 7z &= 4 \\ 9x + 4y + 0z &= 1 \end{aligned}$$

can be represented by the matrix

$$\underbrace{\begin{bmatrix} 1 & 2 & 5 & 8 \\ 1 & 1 & 7 & 4 \\ 9 & 4 & 0 & 1 \end{bmatrix}}_{[A|\vec{b}]} \to \underbrace{\begin{bmatrix} 1 & 0 & 0 & -\frac{135}{73} \\ 0 & 1 & 0 & \frac{322}{73} \\ 0 & 0 & 1 & \frac{15}{73} \end{bmatrix}}_{[U|\vec{y}]}.$$

Thus because $U = I$ we see that x, y, z all have unique solutions.

5.1.5 Observations about U

Here are some observations about U:

- All nonpivotal columns are linear combinations of the previous pivotal columns.
- The solutions to $A\vec{x} = \vec{b}$ is encoded in $\vec{y}$. For example if $U = I$, then $\vec{x} = \vec{y}$.

- Otherwise a particular solutioon will take y_1 of the first pivotal column, y_2 of the second pivotal column, and y_n of the nth pivotal column. Note that the index of y must match with the nth pivotal column. For example if we have the row echelon matrix

$$\underbrace{\left[\begin{array}{cccccccc|c} 0 & \underline{1} & 3 & 0 & 0 & 3 & 0 & -4 \\ 0 & 0 & 0 & \underline{1} & -2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{1} & 2 \end{array}\right]}_{[U|\vec{y}]}$$

then the particular solution will be

$$\vec{x} = \begin{bmatrix} 0 \\ -4 \\ 0 \\ 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}.$$

5.1.6 Solution to $A\vec{x} = \vec{0}$

Now that know that we can use $U\vec{x} = \vec{y}$ to find a particular solution to $A\vec{x} = \vec{b}$, now we want to find the set of solutions to $A\vec{x} = \vec{0}$. How can we do this?

Theorem 4 **Solution to $A\vec{x} = \vec{0}$ using RREF**
Let $A \in \mathbb{R}^{m\times n}$ is a given matrix and suppose $U = \text{RREF}(A)$. Then, for any $\vec{x} \in \mathbb{R}^n$ we have

$$A \cdot \vec{x} = \vec{0} \iff U \cdot \vec{x} = \vec{0}.$$

Proof: We prove both directions of implication.

($\Rightarrow$) First, we show that if $A\vec{x} = \vec{0}$ then $U\vec{x} = \vec{0}$.
Recall that $E \cdot A = U$ where E is a product of elementary matrices. Therefore E is a nonsingular matrix. Therefore

$$\begin{aligned} A\vec{x} &= \vec{0} \\ EA\vec{x} &= E\vec{0} \\ U\vec{x} &= \vec{0}. \end{aligned}$$

($\Leftarrow$) Now, we show that if $U\vec{x} = \vec{0}$ then $A\vec{x} = \vec{0}$.

$$\begin{aligned} U\vec{x} &= \vec{0} \\ EA\vec{x} &= \vec{0} \\ E^{-1}EA\vec{x} &= E^{-1}\vec{0} \\ A\vec{x} &= \vec{0}. \quad \square \end{aligned}$$

5.1.7 Finding Trivial Solutions to $U\vec{x} = \vec{0}$

Looking our matrix from above,

$$\underbrace{\left[\begin{array}{ccccccc|c} 0 & 1 & 3 & 0 & 0 & 3 & 0 & -4 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]}_{[U|\vec{y}]}$$

since there are 4 non-pivotal columns, we know that we will have four linearly independent solutions (one for each non-pivot column), so the solution set of the nontrivial solutions will be

$$\vec{z} = c_1 z_1 + c_2 z_2 + c_3 z_3 + c_4 z_4.$$

Also recall from our observations on U that each non-pivot column is a linear combination of the pivot columns that came before. Thus, since for each "unique" solution z_i we want

$$A\vec{z_i} = 0$$

z_i will take 1 of the nonpivotal columns, and negative of the linear combination of pivotal columns such that. For example the third column of our matrix U above is non-pivotal (it is the second non-pivotal column), and so

$$z_2 = \begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

In this way, the complete solution of $\vec{z}$ is

$$\vec{z} = c_1 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\vec{z}_1} + c_2 \underbrace{\begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\vec{z}_2} + c_3 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{z_3} + c_4 \underbrace{\begin{bmatrix} 0 \\ -3 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{\vec{z}_4}.$$

SECTION 6

Lecture 18: Properties of Determinants

We have seen that if matrix A is invertible then a solution exists to the $A\vec{x} = \vec{b}$ linear systems problem. One way we can determine if A is invertible is with a function called the determinant:

$$\det : \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}, \qquad \begin{cases} \det(A) & \neq 0 \text{ if } A \text{ is invertible} \\ \det(A) & = 0 \text{ if } A \text{ is singular.} \end{cases}$$

Before describing the big formula for the determinant, let's look at the properties of the determinant function.

SUBSECTION 6.1
Definition of the Determinant

Definition 17

Determinant
There exists a unique function D: $(\mathbb{R}^n)^n \to \mathbb{R}$ that is

1. Multilinear: D is linear with respect to each of its arguments.
2. Antisymmetric: Exchanging any two arguments changes its sign.
3. Normalized: $D(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$

The *determinant* of an $n \times n$ matrix $A = [\vec{a}_1, \ldots, \vec{a}_n]$ is

$$\det A = D(\vec{a}_1, , \vec{a}_n).$$

Source: Hubbard

Let's unpack this definition a little bit:

- What are the arguments of D? Well, D takes n column vectors as arguments, each vector of size $\mathbb{R}^{n\times 1}$.
- What does multilinear mean? This is just saying that if one (and only one) of the columns of A is $\vec{a}_i = \alpha\vec{u} + \beta\vec{w}$ then

$$\begin{aligned}\det(A) &= \det([\vec{a}_1, \ldots, \vec{a}_{i-1}, (\alpha\vec{u} + \beta\vec{w}), \vec{a}_{i+1}, \ldots, \vec{a}_n]) \\ &= \alpha \det[\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{u}, \vec{a}_{i+1}, \ldots, \vec{a}_n] \\ &\quad + \beta \det[\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{w}, \vec{a}_{i+1}, \ldots, \vec{a}_n]\end{aligned}$$

  For example

$$\det\left(\begin{bmatrix} 3a+5a' & b \\ 3c+5c' & d \end{bmatrix}\right) = 3\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) + 5\det\left(\begin{bmatrix} a' & b \\ c' & d \end{bmatrix}\right).$$

- Normalized just means that the determinant of the identity matrix is 1.

6.1.1 Proof

Walk through proof from Hubbard here.

SUBSECTION 6.2
Properties of Determinant

1. Determinant of identity matrix is 1.
2. If a row (or column) of A has all zeros, then $\det(A) = 0$.
3. Multiplying matrix A by row-addition elementary matrix does not change determinant. That is $\det(E_2(i, j, x) \cdot A) = \det(A)$.
4. Row exchanges of matrix A flip sign of determinant. That is $\det(E_3(i, j) \cdot A) = -\det(A)$. From this we know that the determinant of a permutation matrix is either 1 or -1. One if the number of exchanges was even, and -1 if the number of exchanges was odd.

5. Scaling a row of matrix A, scales the determinant by the same amount. That is $\det(E_1(i,x)\cdot A) = x\det(A)$, where E_1 is the elementary matrix that scales row i by x.
6. Linearity of each row. That is
$$\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}.$$
Note that we are NOT saying that $\det(A+B) = \det(A) + \det(B)$.
7. If A is upper- or lower-triangular, then $\det(A) = a_{11}\cdot a_{22}\cdots a_{nn}$
8. $\det(A) = \det\left(A^T\right)$
9. $\det(A\cdot B) = \det(A)\cdot\det(B)$
10. If S invertible, then $\det\left(S\cdot A\cdot S^{-1}\right) = \det(A)$
11. If $i \neq k$, then $\det\left(P_{ik}\cdot A\right) = -\det(A)$
12. If $i \neq k$, then $\det\left(S_{ik}(c)\cdot A\right) = \det(A)$
13. For $1 \leq i \leq n, \det\left(D_i(c)\cdot A\right) = c\cdot\det(A)$
14. $\det(cA) = c^n\det(A)$
15. A is invertible if and only if $\det(A) \neq 0$

Source: Jeff's notes and MIT lecture

Subsection 6.3

Permutations and their Signatures

Let's suppose that you have n things, then a permutation is just a rearrangement of the n items. For example if you have the three primary colors { red , green , blue } then there are 6 different ways to reorder the colors. Let's say that you have n things, then you can describe a permutation as follows:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 5 & 1 & 7 & \cdots & 2 & 3 \end{pmatrix}$$

where the top row is your numbers ordered 1 through n and the bottom row are the numbers after the shuffling. This notation is really clear, but it takes a lot of space!

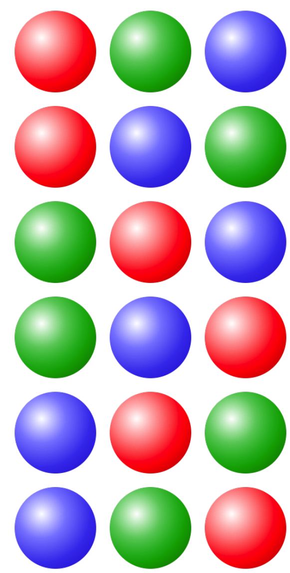


Figure 8. Each row is a different permutation of the three colors.

6.3.1 Cycle Notation

Instead of using this rather cumbersome notation to describe a permutation, cyclic notation can be used instead. To get a better idea of how cycle notation works, let's look at an example.

Example 1

Let's say that we are given the following permuation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}.$$

To convert this to cycle notation perform the following steps:

1. Because there are 5 things, write down the numbers $1 \to 5$: 1 2 3 4 5.

2. Scratch off the number 1 (the rule is to scratch off the number once it appears in the mapping), and note that it maps to 3 (now scratch off the number 3.) We now have the mapping
$$1 \to 3$$
3. We now see that 3 maps to 4, so we cross of 4, and have the mapping
$$1 \to 3 \to 4.$$
4. Note that 4 maps back to 1 giving us the mapping
$$\begin{tikzcd} 1 \arrow[r] & 3 \arrow[r] & 4 \arrow[ll, bend right] \end{tikzcd}.$$
To make this all the more concise, we can simplify this as
$$(1, 3, 4)$$
which just means that 1 maps to 3, 3 maps to 4 and 4 maps back to 1.
5. Note though that we have not covered all the mappings yet, we still have 2 mapping to 5 and 5 mapping to 2. Using the same process as above we thus have
$$(1, 3, 4)(2, 5).$$

Note though that you would get a different representation of the mapping depending upon the number you start with, not to worry though, these are all the same:
$$(1, 3, 4)(2, 5) \Leftrightarrow (4, 1, 3)(5, 2) \Leftrightarrow (3, 4, 1)(2, 5).$$

By convention, for each cycle we start with the smallest number, so we should write the permutation as $(1, 3, 4)(2, 5)$.
Source: Cycle Notation Youtube

Definition 18 **Permutation**
A *permutation* of a set X is a bijective map $f : X \to X$; the word is usually used only when X is finite. The set of permutations of the set $X = \{1, 2, \ldots, n\}$ is denoted Perm_n, which has $n!$ elements. It has the following properties:

1. Composition is associative: if $\sigma_1, \sigma_2, \sigma_3 \in \mathrm{Perm}_n$, then
$$(\sigma_1 \circ \sigma_2) \circ \sigma_3 = \sigma_1 \circ (\sigma_2 \circ \sigma_3)$$
2. There is an identity id for composition: $\sigma \circ \mathrm{id} = \mathrm{id} \circ \sigma = \sigma$.
3. For every $\sigma \in \mathrm{Perm}_n$, the permutation σ^{-1} satisfies
$$\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathrm{id}.$$

Source: Hubbard

Let's unpack this definition

- A bijective map $X \to Y$ just means that the following properties hold:
  i. each element of X must be paired with at least one element of Y,
  ii. no element of X may be paired with more than one element of Y,
  iii. each element of Y must be paired with at least one element of X, and
  iv. no element of Y may be paired with more than one element of X.
- Because it is a map from $X \to X$, this means that we are just going to be rearranging the elements.
- We say that σ_i, where $1 \leq i \leq n!$ is one of the $n!$ permuations.

Definition 19

Signature of Permuation
There exists a unique map

$$\text{sign} : \text{Perm}_n \to \{-1, 1\}$$

called the *signature*, such that

1. $\text{sign}(\sigma_1 \circ \sigma_2) = \text{sgn}(\sigma_1)$
2. $\text{sign}(\sigma_2)$ for all $\sigma_1, \sigma_2 \in \text{Perm}_n$ $\text{sgn}(\tau) = -1$ for all transpositions $\tau \in \text{Perm}_n$

Without getting too into the weeds, all this definition is saying is if there is an even number of switches for the permutation then the signature is even (1), and if there are an odd number of switches then the signature is odd (-1). For example the permutation $(1, 2)$ is odd because you just hade to switch the mappings once.

Subsection 6.4

Permutation Definition of Determinant

Definition 20

Determinant in terms of permutations
Let A be an $n \times n$ matrix with entries $a_{i,j}$. Then

$$\det(A) = \sum_{\sigma \in \text{Perm}_n} \text{sign}(\sigma) a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \ldots a_{n-1,\sigma(n-1)} \cdot a_{n,\sigma(n)}.$$

This is a really confusing definition, what is it saying? All it is saying is that for each term of our summation n elements from our matrix will be multiplied together, and the i, j (row and column position) of those elements will be determined by a given permutation, with i being the number before being shuffled, and j being the number after being shuffled. So in total there will be $n!$ terms and each term will have n elements multiplied together (as well as the signature of the particular permutation we are doing).

6.4.1 A 2 × 2 example

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a square matrix, what is the determinant of this matrix?. Because $n = 2$, there are two total permutations $(1)(2)$ $(1 \to 1,\ 2 \to 2)$ and $(1, 2)$ $(1 \to 2,\ 2 \to 1)$. Making a table of all the permutations, the signature of each permutation and the elements of A being multiplied (determined by the permutation) we have:

$\sigma_1 = (1)(2)$	$+$	$a_{1,1} \cdot a_{2,2} = a \cdot d = ad$
$\sigma_2 = (1, 2)$	$-$	$a_{1,2} \cdot a_{2,1} = b \cdot c = bc$

.

Putting everything together we see that

$$\begin{aligned}\det(A) &= \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \\ &= ad - bc\end{aligned}$$

which aligns with our definition of the determinant we know from calculus.

SECTION 7

Lecture: Eigenvalues and Eigenvectors

Here is where thinking of a matrix encoding a change of basis, and matrix-vector multiplication representing a linear transformation becomes really useful. Recall that a $n \times n$ matrix encodes a change of basis where the 1st column says where $\hat{\imath}$ goes, the second column describes where $\hat{\jmath}$ goes, and so on and so forth, and the result of matrix-vector multiplication just describes where the vector lands after the linear transformation.

What happens though when after the linear transformation, the vector lands on its span? For example, let's say that we have the following matrix-vector multiplication problem:

$$\underbrace{\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 3 \\ 0 \end{bmatrix}}_{\vec{b}}.$$

In this case we see that the $\vec{b}$ remains on the span of $\vec{x}$. That is, the result of the matrix-vector multiplication problem is just scaling the vector:

$$\underbrace{\begin{bmatrix} 3 \\ 0 \end{bmatrix}}_{\vec{b}} = \underbrace{3}_{\lambda} \cdot \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\vec{x}}.$$

The scaling factor λ is an eigenvalue, and the vector being scaled is an eigenvector.

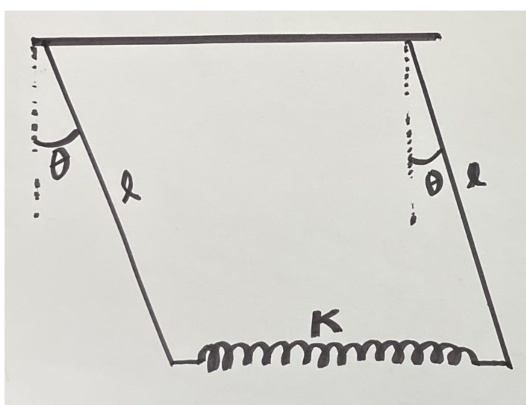
a.png

Figure 9. We see that after the linear transformation the vector $\vec{x}$ remains on its span, scaled by the eigenvalue $\lambda = 3$.

Subsection 7.1

An Applied Problem

Let's say that we two pendulums of equal length connected by a spring, with masses m_1 and m_2:



The question is, can find an equation $\theta(t)$ that tells the position of the pendula at a given time, given some initial conditions (masses of weights and starting position)?

7.1.1 Single Pendulum Problem

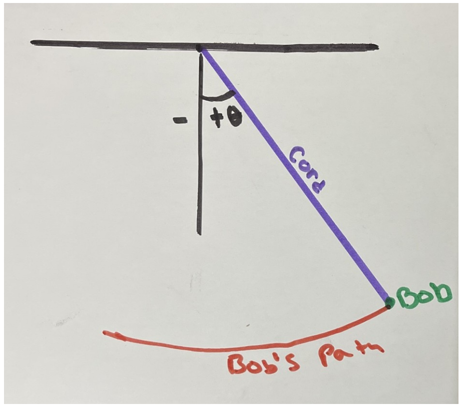


Figure 10. Instead of trying to find the position of the bob at a given time, we will "linearlize" the problem and find the position of the shadow of the bob on the number line.

To create a differential equation that models the motion of our bob, we will use Newtons formula

$$\overrightarrow{\sum F} = m \cdot a.$$

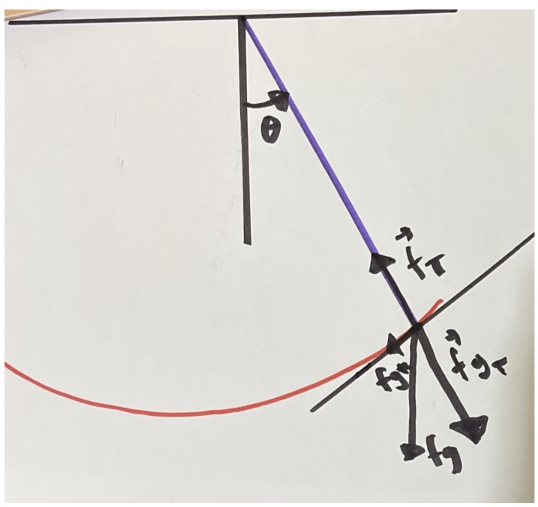


Figure 11. We can decompose the force of gravity. Note tht $\vec{f}_T$ and $\vec{f}_{gT}$ cancel eachother out, so the entire force of the system is $\vec{f}_{G^*_T}$.

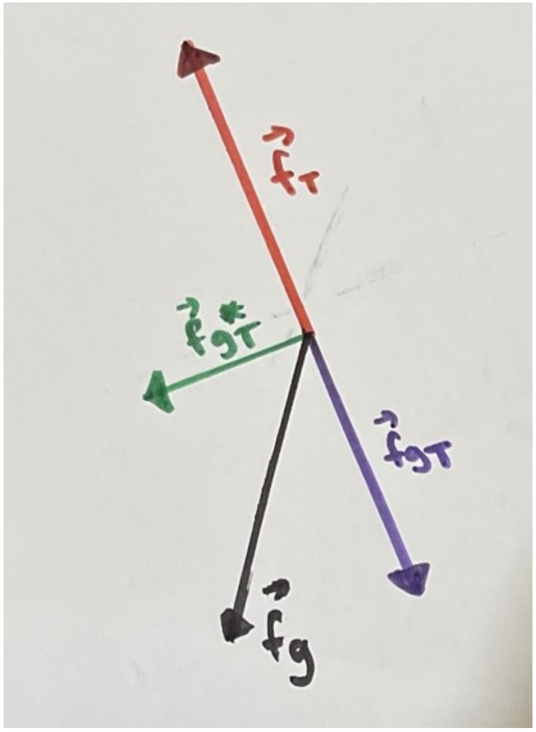


Figure 12. The angle between the force of gravity and the decomposed force of gravity is $-\theta(t)$.

Because of this fact, we can thus say that

$$
\begin{aligned}
|\sin(-\theta(t))| &= \frac{\|\vec{f}_{gT}^{\,*}2}{\|\vec{f}_g\|_2} \\
|\sin(-\theta(t))| \cdot \|\vec{f}_g\|_2 &= \|\vec{f}_{gT}^{\,*}\|_2 \\
|\sin(-\theta(t))| \cdot m\dot{g}
\end{aligned}
$$

Putting everything together we have

$$
\begin{aligned}
\overrightarrow{\sum F} &= m \cdot a(t) \\
-mg\sin(\theta(t)) &= m \cdot l \cdot \ddot{\theta}(t) \\
-\frac{g}{l}\sin(\theta(t)) &= \ddot{\theta}(t) \\
\ddot{\theta}(t) + \frac{g}{l}\sin(\theta(t)) &= 0
\end{aligned}
$$

Note that $a(t)$ is just the double derivative of arc length, $\ddot{s(t)}$. But $s(t) = l\theta(t)$ and so $\ddot{s(t)} = l\ddot{\theta}(t)$.

where g is acceleration due to gravity and l is measured length of pendulum cord. This ODE is unsolvable exactly, but how can we linearlize it? Well, linearlizing the equation we have

$$
\begin{aligned}
\ddot{\theta}(t) + \frac{g}{l}\sin(\theta(t)) &= 0 \\
\ddot{(u)}(t) + \frac{g}{l}u(t) &= 0.
\end{aligned}
$$

7.1.2 Coupled-Pendulum Model

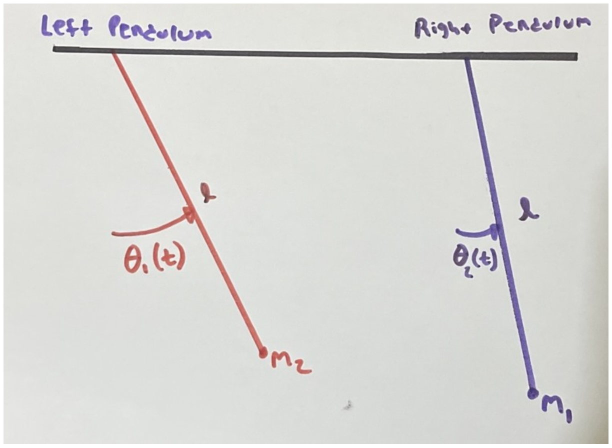


Figure 13. Each pendulum consists of a massless, inextensible cord whose fixed end is attached to a frictionless pivot position on the same stable support structure and whose free end is connected to a bob. Cord length for each cord is identical. The masses are not necessarily the same.

Using all of this information together we can say that for the left pendulum

$$-m_1 \cdot g \cdot \sin(\theta_1(t)) = m_1 \cdot l \cdot \ddot{\theta_1}(t)$$
$$\ddot{\theta}_1(t) + \frac{g}{l}\sin(\theta_1(t)) = 0$$
$$\overrightarrow{\sum F} = m \cdot a(t)$$

and similarily for the right pendulum

$$-m_2 \cdot g \cdot \sin(\theta_2(t)) = m_2 \cdot l \cdot \ddot{\theta_2}(t)$$
$$\ddot{\theta}_2(t) + \frac{g}{l}\sin(\theta_2(t)) = 0$$
$$\overrightarrow{\sum F} = m \cdot a(t).$$

Because we cant solve this ODE exactly we linearlized the problem such that
The force in the spring is

$$\begin{aligned} F_s &= k \cdot e(t) \\ &= k(u_2(t) - u_1(t)). \end{aligned}$$

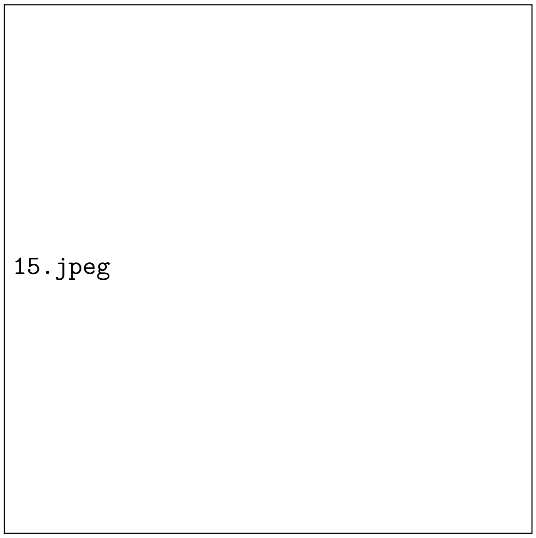


Figure 14. We now have enough information to set up or connected pendulum model.

7.1.3 Free Body Diagram

First what are the forces acting on mass 1? Well, pulling to the left we have the force of pendulum 1 and pulling to the right we have the spring force. Thus

$$\begin{aligned}\underbrace{\sum F_1}_{\substack{\text{Net force acting}\\ \text{on our mass}}} &= f_{p_1}(t) + f_s(t)\\ &= -\frac{m_1 \cdot g}{l} \cdot u_1(t) + k\left(u_2(t) - u_1(t)\right)\\ &= \left(\frac{-m_1 \cdot g}{l} - k\right) u_1(t) + k \cdot u_2(t)\end{aligned}$$

Now looking at the forces on m_2 we have

$$\begin{aligned}\sum F_2 &= f_{p_2}(t) + -f_s(t)\\ &= -\frac{m_2 \cdot g}{l} u_2(t) - k(u_2(t) - u_1(t))\\ &= +k \cdot u_1(t) + \left(-k - \frac{m_2 \cdot g}{l}\right) u_2(t).\end{aligned}$$

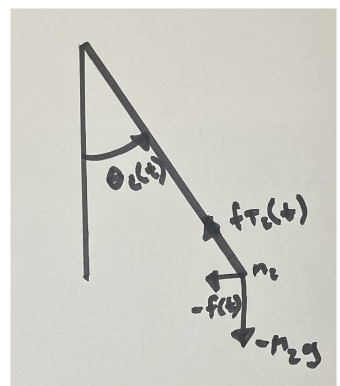


Figure 15. Free body diagram for m_2.

7.1.4 Newton's Second Law

Recalling Newton's second law, we can describe our problem in terms of matrix vector multiplication:

$$
\begin{aligned}
\begin{bmatrix} \sum F_1 \\ \sum F_2 \end{bmatrix} &= \begin{bmatrix} m_1 \cdot \ddot{u}_1(t) \\ m_2 \cdot \ddot{u}_2(t) \end{bmatrix} \\
&= \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} \\
\begin{bmatrix} \left(-\frac{m_1 \cdot g}{l} - k\right) u_1(t) + k u_2(t) \\ k u_1(t) + \left(-k - \frac{m_2 \cdot g}{l}\right) u_2(t) \end{bmatrix} &= \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} \\
\begin{bmatrix} \frac{m_1 \cdot g}{l} - k \\ k \\ k \\ -k - \frac{m_2 \cdot g}{l} \end{bmatrix} \begin{bmatrix} u_t(t) \\ u_2(t) \end{bmatrix} &= \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} \\
-\begin{bmatrix} m_1 \cdot gl + k \\ -k \\ -k \\ \frac{k + m_2 \cdot g}{l} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} &= \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} \\
-K \cdot \vec{u}(t) &= M \ddot{u}(t).
\end{aligned}
$$

where $\vec{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$.

7.1.5 Forming Eigenvalue Problem

To form a standard eigenvalue problem (and not a general one) we will assume that $m_1 = m_2$. Doing this we have

$$
\begin{aligned}
\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} &= -\begin{bmatrix} \frac{m \cdot g}{l} + k & -k \\ -k & k + \frac{mg}{l} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \\
\begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} &= -\begin{bmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} \frac{m-g}{l} + k & -k \\ -k & k + \frac{m \cdot g}{l} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \\
\begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} &= -\begin{bmatrix} \frac{g}{l} + \frac{k}{m} & -\frac{k}{m} \\ -\frac{k}{m} & \frac{k}{m} + \frac{g}{l} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \\
&= \ddot{u}(t) \qquad = -A \cdot \vec{u}(t)
\end{aligned}
$$

where $A = M^{-1} \cdot k$. That is, when trying to figure out the motion of the pendulum, that problem is equivalent to solving the above ODE. We will try and match displacement behavior within the time interval

$$t \in [t_s, t_e].$$

Within our observation, we will choose a reference time $t_0 \in [t_s, t_e]$, a given position and a given velocity. We can guess that

$$
\begin{aligned}
\vec{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} &= \begin{bmatrix} v_1 \ \cos\left(\omega\left(t - t_0\right)\right) \\ v_2 \ \cos\left(\omega\left(t - t_0\right)\right) \end{bmatrix} \\
&= \begin{bmatrix} \cos\left(\omega\left(t - t_0\right) \cdot v_1\right. \\ \cos\left(\omega\left(t - t_0\right) \cdot v_2\right. \end{bmatrix} \\
&= \cos\left(\omega\left(t - t_0\right)\right) \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}
\end{aligned}
$$

Therefore $\vec{u}(t) = \cos(w(t - t_0)) \cdot \vec{v}$, where $\vec{v}$ is a constant vector. This would mean that

$$\begin{aligned}
\ddot{u}(t) &= -A\vec{u}(t) \\
\frac{d^2}{dt^2}[\vec{u}(t)] &= -A \cdot \vec{u}(t) \\
\begin{bmatrix} \frac{d^2}{dt^2}[u_1(t)] \\ \frac{d^2}{dt^2}[u_2(t)] \end{bmatrix} &= -\begin{bmatrix} \frac{g}{2} + \frac{k}{m} & -\frac{k}{m} \\ -\frac{k}{m} & \frac{k}{m} + \frac{g}{2} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \\
\begin{bmatrix} -\omega^2 \cdot v_1 \cdot \cos(\omega(t - t_0)) \\ -\omega^2 \cdot v_2 \cdot \cos(\omega \cdot (t - t_0)) \end{bmatrix} &= -A \cdot \left(\cos(\omega(t - t_0)) \cdot \vec{V}\right) \\
-\omega^2 \cdot \cos(\omega(t - t_0)) \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} &= -\cos(\omega(t - t_0)) \cdot A \cdot \vec{V} \\
\cos(\omega(t - t_0)) \cdot \omega^2 \cdot \vec{v} &= \cos(\omega(t - t_0)) \cdot A \cdot \vec{v} \\
\cos(\omega(t - t_0)) A \cdot \vec{v} - \cos(\omega(t - t_0)) \cdot \omega^2 \cdot \vec{v} &= \overrightarrow{0} \\
\cos(\omega(t - t_0)) \cdot \left(A\vec{v} - \omega^2 \cdot \vec{v}\right) &= \overrightarrow{0}.
\end{aligned}$$

This must hold for all t, and since $\vec{v} \neq 0$. We are thus looking for nonzero vector $\vec{v}$ such that

$$\begin{aligned}
A\vec{v} - \omega^2 \cdot \vec{v} &= 0 \\
A\vec{v} &= \omega^2 \vec{v} \\
A\vec{v} &= \lambda \vec{v} && \textit{For } \lambda = \omega^2.
\end{aligned}$$

Theorem 5 If $\vec{V} \in \mathbb{R}^2$ is an eigenvector of the matrix $A = M^{-1} \cdot K$ with corresponding eigenvalue $\lambda = \omega^2$ s.t.

$$A\vec{v} = \lambda \vec{v}$$

then both vector-valued functions

$$\begin{aligned}
\vec{u}(t) &= \cos(\omega(t - t_0)) \cdot \vec{v} \\
\vec{u}(t) &= \sin(\omega(t - t_0)) \cdot \vec{v}
\end{aligned}$$

solve the 2nd-order system: $\ddot{u}(t) = -A\vec{u}(t)$.

We can use what we have learned in linear algebra to solve:

$$\begin{aligned}
A\vec{v} &= \lambda \vec{v} \\
A_{\vec{v}} - \lambda \vec{v} &= \vec{0} \\
A\vec{v} - \lambda I_2 \vec{v} &= \vec{0} \\
(A - \lambda I_2)\vec{v} &= \vec{0}.
\end{aligned}$$

When does a non-zero vector that sends the matrix to zero? This is only possible if the matrix is singular, which is only true if $\det(A - \lambda I_2) = 0$. Computing the determinant such that it equals zero we have

$$
\begin{aligned}
0 &= \det\left(A - \lambda I_2\right) \\
&= \left[\left(\frac{g}{l} + \frac{k}{m}\right) - \lambda\right]^2 - \left[-\frac{k}{m}\right]^2 \\
&= \left[\left(\frac{g}{2} + \frac{k}{m}\right) - \lambda\right]^2 - \left[\frac{k}{m}\right]^2 \\
\left[\left(\frac{g}{l} + \frac{k}{m}\right) - \lambda\right]^2 &= \left[\frac{k}{m}\right]^2 \qquad voffset \\
\sqrt[2]{\left[\left(\frac{g}{l} + \frac{k}{m}\right) - \lambda\right]^2} &= \sqrt[2]{\left[\frac{k}{m}\right]^2} \\
\left|\frac{g}{l} + \frac{k}{m} - \lambda\right| &= \frac{k}{m} \\
\lambda - \left(\frac{g}{l} + \frac{k}{m}\right) &= \pm\frac{k}{m}.
\end{aligned}
$$

There are thus two solutions to our characteristic equation (two distinct eigenvalues).

- Eigenvalue 1: First Natural Frequency

$$
\begin{aligned}
\lambda_1 = \left(\frac{g}{l} + \frac{k}{m}\right) - \frac{k}{m} &= \frac{g}{l} \\
&= \omega_1^2 \\
&= \sqrt{\frac{g}{l}}.
\end{aligned}
$$

  This is the same frequency as the simple pendulum.

- Eigenvalue 2: Second Natural Frequency

$$
\begin{aligned}
\lambda_2 &= \left(\frac{g}{l} + \frac{k}{m}\right) + \frac{k}{m} \\
&= \frac{9}{l} + \frac{2k}{m} \\
&= \omega_2^2 \\
&= \sqrt{\frac{g}{l} + \frac{2k}{m}}
\end{aligned}
$$

Recall that

$$A\vec{v}_1 = \lambda_1\vec{v}_1$$

and so

$$\lambda_1 = \frac{g}{l} \quad \text{and} \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with $\omega_1^2 = \lambda_1$.

The Eigenvector 2. We want to find the nonzero vector $\vec{v}_2$ such that

$$A\vec{v}_2 = \lambda_2\vec{v}_2.$$

Therefore

$$(A - \nabla_2 I_2)\vec{v}_2 = \vec{0}$$

$$\begin{bmatrix} \left(\frac{g}{l} + \frac{k}{m}\right) - \left(\frac{g}{l} + \frac{2k}{m}\right) & -\frac{k}{m} \\ -\frac{k}{m} & \left(\frac{k}{m} + \frac{g}{l}\right) - \left(\frac{g}{l} + \frac{2k}{m}\right) \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad voffset$$

$$\begin{bmatrix} -\frac{k}{m} & -\frac{k}{m} \\ -\frac{k}{m} & -\frac{k}{m} \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad voffset$$

We therefore have

$$\vec{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

We therefore have

$$A\vec{v}_2 = \lambda_2 \vec{v}_2$$

where

$$\lambda_2 = \frac{g}{l} + \frac{2k}{m} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

with

$$\omega_2^2 = \lambda_2 \Leftrightarrow \omega_2 = \sqrt{\frac{g}{l} + \frac{2k}{m}}.$$

7.1.6 Diaganalization

We can group the eigenvalue-eigenvector pairs in the following way:

$$\begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \vec{v}_1 & \lambda_2 \vec{v}_2 \end{bmatrix}$$

$$A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$[A] \cdot [V] = [\mathrm{V}] \cdot [\Lambda]$$

$$\begin{bmatrix} \frac{g}{l} + \frac{k}{m} & \frac{-k}{m} \\ \frac{-k}{m} & \frac{k}{m} + \frac{g}{l} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{g}{2} & 0 \\ 0 & \frac{g}{2} + \frac{2k}{m} \end{bmatrix}$$