Linear Algebra
Week 4

G-07

17 X 2024

1 Gauss Elimination

We can represent linear systems of equations in form of $A\mathbf{x} = \mathbf{b}$ using matrix vector multiplication. But how about solving the LSE? Our first algorithm is *Gauss Elimination*, which actually is the classical algorithm for solving linear systems of equations. It consists of two steps: elimination and back substitution. For this week we assume that our coefficient matrix is $m \times m$, which means that we have a linear system of equations in m equations and m variables.

If you find a step difficult to understand, go to the subsection *1.3 Solving $A\boldsymbol{x} = \boldsymbol{b}$ using Gaussian Elimination* to see the theory in practice.

1.1 Elimination

We want to generate an upper triangular matrix U using our initial matrix A and the vector $\mathbf{b}$. We want to transform the equation $A\mathbf{x} = \mathbf{b}$ into another equation $U\mathbf{x} = \mathbf{c}$ where U is an upper triangular matrix and $\mathbf{x}$, the solutions are the same for both systems. In other words we want to create *equivalent* systems.

It is going to become obvious why we want an upper triangular matrix when we handle back substitution.

There are two main row operations to perform the elimination:

- row subtraction
- row exchange

These operations do not change the solution of the system when applied on both A and $\mathbf{b}$. This is why we usually concatenate A and $\mathbf{b}$ when we perform these elimination steps as follows:

$$\left[\begin{array}{c|c} & \\ A & \mathbf{b} \\ & \end{array}\right]$$

Both operations can be represented as linear transformations!
In the following there are the matrices that represent these linear transformations in the general case:

$$E_{ij} = \begin{bmatrix} \ddots & & & \\ & 1 & & \\ & & \ddots & \\ & -c & & 1 & \\ & & & & \ddots \end{bmatrix} \begin{array}{l} \\ \leftarrow j \\ \\ \leftarrow i \\ \\ \end{array} \qquad \begin{array}{cc} \uparrow & \uparrow \\ j & i \end{array}$$

(a) Elimination

$$P_{jk} = \begin{bmatrix} \ddots & & & \\ & 0 & & 1 & \\ & & \ddots & & \\ & 1 & & 0 & \\ & & & & \ddots \end{bmatrix} \begin{array}{l} \\ \leftarrow j \\ \\ \leftarrow k \\ \\ \end{array} \qquad \begin{array}{cc} \uparrow & \uparrow \\ j & k \end{array}$$

(b) Permutation

Here E_{ij} represents the elimination matrix that is supposed to create a zero entry in row i and column j and the permutation matrix P_{jk} changes row j with row k.

When do we use the row operations?
We try to generate an upper triangular matrix with nonzero elements on the diagonal. So if we have a nonzero element under the diagonal, we subtract one column from the other to get rid of it. We exchange rows when we get a zero element on the diagonal,*i.e.* when we have a pivot that is equal to 0.

The diagonal entry of the current column is called the **pivot**.

After completing the elimination we are left with the equation $U\mathbf{x} = \mathbf{b}$ that has the same solutions as our initial equation $A\mathbf{x} = \mathbf{b}$. Then we continue with back substitution.

1.2 Back Substitution

After we get to an upper triangular matrix, we have our equation suitable for back substitution. This step is not different than what you would do in high school to solve a linear system of equations. Let's see it in the 3×3 example from the lecture notes:

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 19 \\ 17 \\ 14 \end{bmatrix}$$

And the back substitution in the direction of the arrow on the right:

equation	before substitution	after substitution	solution	↑
1	$2x_1 + 3x_2 + 4x_3 = 19$	$2x_1 + 11 = 19$	$x_1 = 4$	
2	$5x_2 + 6x_3 = 17$	$5x_2 + 12 = 17$	$x_2 = 1$	
3	$7x_3 = 14$		$x_3 = 2$	

Using the variables, for which we already have calculated the value, we can calculate the values of other variables. A way to formalize this is:

$$x_i = \frac{b_i - \sum_{j=i+1}^{m} a_{ij}x_j}{a_{ii}}$$

Since we know all x_j here, we can calculate x_i.

1.3 Solving $A\mathbf{x} = \mathbf{b}$ using Gauss Elimination

We are given the system of linear equations represented as:

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

We wish to solve the system $A\mathbf{x} = \mathbf{b}$. First, we will perform Gaussian elimination by applying elimination matrices and a permutation matrix.

- Step 1: Apply a Permutation Matrix to Swap Rows

  Since the first pivot is zero, we swap row 1 with row 2 using a permutation matrix P:

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  After applying the permutation, the augmented matrix becomes:

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

- Step 2: First Elimination Matrix E_{31}

  We eliminate the entry below the first pivot (2 in the $(1,1)$ position) by subtracting half of row 1 from row 3. The elimination matrix E_{31} is:

$$E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$$

  After applying E_{31}, the augmented matrix becomes:

$$\begin{bmatrix} 2 & -3 & 4 & \mid & -2 \\ 0 & 2 & 1 & \mid & 1 \\ 0 & \frac{5}{2} & -4 & \mid & 1 \end{bmatrix}$$

- Step 3: Second Elimination Matrix E_{32}

  We now eliminate the entry below the second pivot (2 in the $(2,2)$ position) by subtracting $\frac{5}{4}$ of row 2 from row 3. The elimination matrix E_{32} is:

$$E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{5}{4} & 1 \end{bmatrix}$$

After applying E_{32}, the augmented matrix becomes:

$$\begin{bmatrix} 2 & -3 & 4 & | & -2 \\ 0 & 2 & 1 & | & 1 \\ 0 & 0 & -\frac{11}{4} & | & \frac{9}{4} \end{bmatrix}$$

Final Upper Triangular Form

Thus, we have reduced the system to an upper triangular form. The final system is:

$$U\mathbf{x} = \mathbf{c}$$

where:

$$U = \begin{bmatrix} 2 & -3 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & -\frac{11}{4} \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} -2 \\ 1 \\ \frac{9}{4} \end{bmatrix}$$

We can now solve for $\mathbf{x}$ by back-substitution.

Question 4.1 Perform the back substitution.

1.4 Success and Failure

There are 3 important points in this section that further explain how Gauss elimination works and when it fails.

Lemma 3.3. *Let $A\mathbf{x} = \mathbf{b}$ be a system of m linear equations in n variables, and let $M \in \mathbb{R}^{m\times m}$ be a row operation matrix. Let $A' = MA$ and $\mathbf{b}' = M\mathbf{b}$ be the result of applying the row operation to both A and $\mathbf{b}$. Then the two systems $A\mathbf{x} = \mathbf{b}$ and $A'\mathbf{x} = \mathbf{b}'$ have the same solutions.*

Corollary 3.4. *Let A be an $m \times n$ matrix, let $M \in \mathbb{R}^{m \times m}$ be a row operation matrix, and let $A' = MA$ be the result of applying the row operation to A. Then A has linearly independent columns if and only if A' has linearly independent columns.*

Theorem 3.5. *Let $A\mathbf{x} = \mathbf{b}$ be a system of m linear equations in m variables. The following two statements are eqivalent.*

(i) Gauss elimination as in Table 3.6 succeeds.

(ii) The columns of A are linearly independent.

I have nothing to add to these facts, and their proofs are documented formally in the lecture notes. However, as a small digression from the lecture notes, I can provide some intuition about what it exactly means when we fail in the pdf named: *What does failure actually mean?*

1.5 Runtime

It is always good to know the runtime of an algorithm you use. Here it is for Gauss Elimination, taken from the lecture notes. You can find the full runtime analysis in the lecture notes.

Theorem 3.6. *Let $A\mathbf{x} = \mathbf{b}$ be a system of m linear equations in m variables, $m \geq 1$. Gauss elimination with back substitution solves $A\mathbf{x} = \mathbf{b}$ (or gives up) with at most*

$$g(m) = \frac{2}{3}m^3 + \frac{3}{2}m^2 - \frac{7}{6}m$$

arithmetic steps and therefore in time $O(m^3)$.

2 Inverse Matrices

Until now we have seen how to add matrices, how to multiply them, but not how to "divide" by a matrix. If you see a matrix as a linear transformation, this division operation would correspond to undoing the effect of the linear transformation. Did your matrix stretch the space and doubled every distance in x_1 direction? Then the inverse would shrink the space in the same direction so that everything gets back to "normal", or in other words everything gets back to the way as they were before you applied any transformations.

2.1 Definition and Properties

Definition 3.7 (Invertible matrix) Let M be an $m \times m$ matrix. M is called invertible if there exists an $m \times m$ matrix M^{-1} (called the inverse of M) such that

$$MM^{-1} = M^{-1}M = I$$

Note that the matrix M is given as a square matrix and hence the inverse is also a square matrix. If this was not the case and the dimensions do not match, then it would be likely that one of the matrix products above would not even be defined, let alone the inverse matrix. But in this case where we have an $m \times m$ square matrix and when we know the inverse exists, one of these expressions imply the other. In other words if M^{-1} is the right inverse of M then it is also the left inverse. (You are going to prove this fact in this weeks exercises.)

How to calculate the inverse?

Case 1×1**.**

$$M = \begin{bmatrix} a \end{bmatrix} \quad \Rightarrow \quad M^{-1} = \begin{bmatrix} \frac{1}{a} \end{bmatrix} \quad \text{(if } a \neq 0\text{)}.$$

Hence, the inverse exists unless $a = 0$.

Case 2×2**.**

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \Rightarrow \quad M^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{(if } ad - bc \neq 0\text{)}.$$

The inverse for 2×2 matrices exist, if and only if when $ad - bc \neq 0$. This is equivalent to say that the inverse of the 2×2 matrix exists, if and only if the columns of it are linearly independent. (Think about this!).

This formula might feel like it came out of nowhere, but these two examples above are only special cases of a general rule to calculate the inverse of a matrix. The formula utilizes some linear algebra tools we have not acquired yet, like determinants and cofactors, but don't worry, we are going to get there.

Some important properties of inverse matrices are given as 3 lemmas in the lecture notes, which are also provided here with proofs.

If a matrix has an inverse, it is unique

Lemma 3.8. *Let M be an $m \times m$ matrix with two inverses A and B. Then $A = B$.*

Proof. Using associativity of matrix multiplication (Lemma 2.22) as well as Corollary 2.20, we compute

$$A = IA = (BM)A = B(MA) = BI = B.$$

□

The inverse of a product of two invertible matrices

Lemma 3.9. *Let A and B be invertible $m \times m$ matrices. Then AB is also invertible, and*

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

Recall that AB is the matrix of the linear transformation $T_{AB}(\mathbf{x}) = T_A(T_B(\mathbf{x}))$ (first apply T_B, then T_A); see Lemma 2.30. To undo this, we need to apply the inverse transformations in reverse order: $T_{B^{-1}}(T_{A^{-1}}(\mathbf{x})) = T_{B^{-1}A^{-1}}(\mathbf{x})$ (first undo T_A, then undo T_B). You can already consider this a proof sketch of the lemma, but there is also a more direct proof not involving linear transformations.

Proof.

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

and

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

□

This one also applies for more than two matrices. You can see it by using associativity of matrix multiplication.

Inversion and Transposition

Lemma 3.10. *Let A be an invertible $m \times m$ matrix. Then the transpose matrix $A^\top$ is also invertible, and*

$$\left(A^\top\right)^{-1} = \left(A^{-1}\right)^\top.$$

Proof. We need to check that

$$\underbrace{A^\top \left(A^{-1}\right)^\top}_{(A^{-1}A)^\top} = \underbrace{\left(A^{-1}\right)^\top A^\top}_{(AA^{-1})^\top} = I,$$

and this is true since we can invoke Lemma 2.19 to pull the transpositions out (below the curly braces) and then use that A and A^{-1} are inverse to each other, alongside with $I^\top = I$. □

2.2 The Inverse Theorem

The inverse theorem is your friend, embrace it, understand it, use it. Here I provide it without the proof, which you can read in the lecture notes. I strongly recommend reading it.

Theorem 3.11 (Inverse Theorem). *Let A be an $m \times m$ matrix. The following statements are equivalent.*

(i) A is invertible.

(ii) For every $\mathbf{b} \in \mathbb{R}^m$, $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}$.

(iii) The columns of A are linearly independent.

3 Hints

1. Solved In Class

2. Try to represent $AA^{-1} = I$ as three separate systems of equations. See task 3a). Be careful with division. You can divide by a variable μ that might be zero **with the constraint that** $\mu \neq 0$ and afterwards you must specify what happens when that variable is zero. In other words make a small case distinction.

3. **a)** The solutions to the system can be read off without elimination. **b)** almost the same as 2nd exercise **c)** no hints **d)** Combine your results from the previous subtasks.

4. **a)** use induction over k **b)** Go to the assignment from week 2 exercise 2c **c)** This is a one liner think clever **d)** you have already seen a type of matrix in the course, which can be a self inverse, find it. (you are of course free to come up with another one) **e)** If you can modify the one from the previous subtask, go for it.

5. In this exercise, at all steps do not forget that you are given $AB = I$. Use that fact. **a)** columns of B are linearly independent if $\mathbf{0}$ is the only solution to $B\mathbf{x} = \mathbf{0}$ Prove if $B\mathbf{x} = \mathbf{0}$ then $\mathbf{x}$ has to be $\mathbf{0}$. **b)** same thing you did in the previous subtask for B, now do it for A and note that there exists such an x so that B$\mathbf{x} = \mathbf{y}$ for which we have $A\mathbf{y} = \mathbf{0}$ (why?) **c)** You have proven $A\mathbf{y} = \mathbf{0}$ in b) use this here.

6. (Challenging) **a)** use formula **b)** You have to prove two directions. Drawing a sketch or writing down such a matrix might help. For the proof of one direction, the proof for Theorem 3.5. $\neg(i) \Rightarrow \neg(ii)$ might inspire you. **c)** Assume for a contradiction that the inverse of a lower triangular matrix is not lower triangular. What is the issue here? Argue formally for the general case. **d)** No hints. You can still do this one without actually solving the first three subtasks.

mkilic

♠