Linear Algebra
Week 5

G-17

23 X 2025

1 Linear Systems of Equations (LSE's)

We are going to handle equations of type A$\mathbf{x} = \mathbf{b}$ in this chapter. Let's see how linear systems of equations can be represented as a matrix and a vector. Let our LSE be as in the following:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

This suspiciously looks similar to a matrix vector multiplication! If it is not clear to you how, notice that each row i of coefficients $a_{i1} \dots a_{in}$ gets multiplied with our unknown variables x_i. Seeing the matrix-vector picture might help:

$$A\mathbf{x} = \mathbf{b} : \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{A,\, m\times n} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}\in\mathbb{R}^n} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}\in\mathbb{R}^m}.$$

2 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 classic 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. This is the *nice* case, we will deal with different number of equations and variables later.

If you find a step difficult to understand, go to the subsection below *2.3 Example: Solving $A\boldsymbol{x} = \boldsymbol{b}$ using Gaussian Elimination* to see the theory in practice.

2.1 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:

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

(a) Elimination

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

(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} swaps row j with row k. To better convince yourself to this you can think of the matrix multiplication $E_{ij}A$ as linear combinations of the rows of A using the row vectors of E_{ij} as the coefficients. The i-th row of $E_{ij}A$ is $-c$ times the jth row of A plus 1 times the ith row of A.

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 $A\mathbf{x} = \mathbf{b}$. We are going to prove this but first let's continue solving the LSE with back substitution:

2.2 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	$2x_1 + 3x_2 + 4x_3 = 19$	$2x_1 + 11 = 19$	$x_1 = 4$	$\uparrow$
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.

2.3 Example: 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:

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

- 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 matrix becomes:

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

- 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:

$$\underbrace{\begin{bmatrix} 2 & -3 & 4 & | & -2 \\ 0 & 2 & 1 & | & 1 \\ 0 & 0 & -\frac{11}{4} & | & \frac{9}{4} \end{bmatrix}}_{U \quad | \quad \mathbf{c}}$$

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.
From the third equation:

$$-\tfrac{11}{4}x_3 = \tfrac{9}{4} \quad \Rightarrow \quad x_3 = \frac{\frac{9}{4}}{-\frac{11}{4}} = -\tfrac{9}{11}.$$

From the second equation using $x_3 = -\frac{9}{11}$:

$$2x_2 + x_3 = 1 \quad \Rightarrow \quad 2x_2 = 1 - x_3 = 1 + \tfrac{9}{11} = \tfrac{20}{11} \quad \Rightarrow \quad x_2 = \tfrac{10}{11}.$$

From the first equation using $x_2 = \frac{10}{11}$ and $x_3 = -\frac{9}{11}$:

$$2x_1 - 3x_2 + 4x_3 = -2 \quad \Rightarrow \quad 2x_1 - 3\left(\tfrac{10}{11}\right) + 4\left(-\tfrac{9}{11}\right) = -2.$$

$$2x_1 - \tfrac{30}{11} - \tfrac{36}{11} = -2 \quad \Rightarrow \quad 2x_1 - \tfrac{66}{11} = -2 \quad \Rightarrow \quad 2x_1 - 6 = -2 \quad \Rightarrow \quad x_1 = 2.$$

And we are done! The vector $\mathbf{x} \in \mathbb{R}^3$ that satisies $A\mathbf{x} = \mathbf{b}$ is:

$$\mathbf{x} = \begin{pmatrix} 2 \\ \frac{10}{11} \\ -\frac{9}{11} \end{pmatrix}.$$

We found a solution! However this is not always the case. Depending on the particular LSE there might be infinitely many solutions or no solutions at all and our algorithm might fail in some cases. To be more precise, we get stuck if we can't make the pivot nonzero using row exchanges. You can see the pdf on my website *What does failure actually mean in Gaussian Elimination?* to gain a little more insight on what might happen when the algorithm fails.

2.4 Row Operations: the general picture

There are important lemmas that formalize our intuition about the algorithm and actually prove that it solves the LSE. The most important one is:

Lemma 3.2 (Invariance of solutions). Let A be an $m \times n$ matrix and M an invertible $m \times m$ matrix. Then the two systems $A\mathbf{x} = \mathbf{b}$ and $MA\mathbf{x} = M\mathbf{b}$ have the same solutions $\mathbf{x}$.

Lemma 3.2 justifies that Gauss elimination with back substitution finds the solution of the actual LSE and does not change it with row subtractions and permutations. Note that both row subtractions and permutations are repreented with invertible matrices. To invert them simply add the row back or swap the two rows back!

Lemma 3.3 (Invariance of the nullspace). Let A be an $m \times n$ matrix and M an invertible $m \times m$ matrix. Then A and MA have the same nullspace, $\mathcal{N}(A) = \mathcal{N}(MA)$.

Lemma 3.4 (Invariance of linear independence). Let A be an $m \times n$ matrix and M an invertible $m \times m$ matrix. Then the following is true: A has linearly independent columns if and only if MA has linearly independent columns.

You can use ***Lemma 3.3*** to prove ***Lemma 3.4***

Lemma 3.5 (Invariance of the row space). Let A be an $m \times n$ matrix and M an invertible $m \times m$ matrix. Then A and MA have the same row space, $\mathcal{R}(A) = \mathcal{R}(MA)$.

This has a beautiful proof in the lecture notes and I recommend to read all proofs, but to convince you to this fact (***Lemma 3.5.***) you can think of the gauss eliminations as creating new linear combinations of the rows of A. Because the matrix multiplication MA can be considered as the linear combination of the rows of A using the rows of M as coefficients. The rows of M are themselves linearly independent (because M is invertible), so the linear combinations of the rows of A are going to have the same span they had before the linear combination.[1]

One must be careful that the same property does **NOT** hold for the column space. In general $C(A)$ is not equal to $C(MA)$. You can find a counterexample in the lecture notes. We can however come to a conclusion about the rank of both matrices.

Lemma 3.6 (Invariance of independent column indices and rank). Let A be an $m \times n$ matrix, M an invertible $m \times m$ matrix. Then the following is true for all $j \in [n]$: column j of A is independent if and only if column j of MA is independent. In particular, A and MA have the same number of independent columns and therefore also the same rank.

Finally, we can say when the elimination fails and when it succeeds. This is the case when the columns of A are linearly independent. This makes sense since m linearly independent vectors can span the whole $\mathbb{R}^m$ so there exists a unique $\mathbf{x} \in \mathbb{R}^m$ for every $\mathbf{b} \in \mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{b}$, given that the columns of A are linearly independent. The following theorem formalizes this. You can find a beautiful proof in the lecture notes.

Theorem 3.8 (Solving $A\mathbf{x} = \mathbf{b}$ with Gauss elimination). Let $A\mathbf{x} = \mathbf{b}$ be a system of m linear equations in m variables. If A has linearly independent columns, then Gauss elimination (Algorithm 2) with back substitution (Algorithm 1) computes the unique solution $\mathbf{x}$ of the system. If A has linearly dependent columns, then Gauss elimination fails.

[1] Keep in mind that this is not a formal proof.

2.5 Runtime

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

For m equations and m unknowns:

- Gauss Elimination: $O(m^3)$
- Back Substitution: $O(m^2)$
- Gauss Elimination with Back Substitution: $O(m^3)$
- Calculating the inverse of an invertible $m \times m$ matrix as in Algorithm 3 from the lecture notes: $O(m^3)$

To see how one can calculate the inverse using Gauss Elimination and Back Substitution you can see the corresponding document on my website.

3 Hints

1. Solved In Class

2. A is invertible if and only if $A\mathbf{x} \neq \mathbf{0}$ for all $\mathbf{x} \neq 0$. Use it and show that none of the entries of $A\mathbf{x}$ is equal to 0.

3. This one is very similar to the in class exercise. Make a linear system of equations out of the given data points. Remember, the variables you want to find are a, b, c.

4. Try to represent $AA^{-1} = I$ as three separate systems of equations. 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.

5. You can first guess the inverse and then verify that $AA^{-1} = I$ by direct computation or calculate the inverse using Gauss elimination with back substitution. **b)** Assume that the pattern from A holds in general. Try to generalize it by writing the columns of the inverse using unit vectors. Then show that the matrix you described is actually the inverse.

6. Try to write the equations using matrices as in $W = VM$.

mkilic

♠