Lineare Algebra

LGS - Lineare Gleichung System - linear system of equations

Vektoren

Lineare kombination - Summe von skalierten Vektoren
Linearly dependent - When two vectors can be expressed as a linear combination of the other and thus doesn't add any information to a LGS. *Basis* - the set of linearly independent vectors $e_1...e_n$ that span all of space R^n

Vektoren werden immer als Spalten in diesem Kurs gezeichnet.

Matrix multiplication comes from the motivation for an efficient way of representing linear combinations / transformations of space.

Geometry of an LGS (Beispiel 1.1.0.8)

An LGS can be viewed geometrically (2D/3D) in multiple different ways:

1. A linear combination of vectors (the columns of the matrix), where we are solving for the set of scalars where the superposition of the vectors is equal to the RHS. The columns of the matrix can be viewed as basis vectors of a custom coordinate system, in which we need to find the equivalent of the RHS vector.
2. Alternatively it can be viewed as a set of line / plane equations (where each row is the normal vector to the plane, unsure if the coefficients are meaningful in ax + by=c) and solutions are points / lines of intersection.
3. The LHS can also be viewed (usually in 3B1B videos) as a linear transformation of space, where the columns of the matrix are where the basis vector of each dimension lands after the transformation. The solution is therefore the vector which after being transformed results in the RHS vector.

NOTE: I will mostly think in terms of the *linear transformation of space* intuition, because the others are not very meaningful when considering inverse matrices.

Superposition:

In this example, one of the LHS vectors is a linear combination of the other two. This results in the LGS only being able to express vectors in a single plane rather than the entire 3D space (it doesn't contain a 3rd base component).

Infinite solutions - if the RHS vector lays in the plane expressed by a_{1-3}, any point in the positive / negative direction of the solution vector lays in the plane.
No solutions - the vector does not lay perfectly on the plane, the LHS vectors lack a component (not necessarily base unit vector) in its direction.

Line / Plane equations:

The solution is the point at which the lines / planes represented by the horizontal equations intersect. There are many possible arrangements which we can visualize, especially in $3D$ space.

Unique solution - Common point of intersection of n non parallel lines / planes.
Infinite solutions - Sheaf of planes or if all lines are the same.
No solution - Not all lines / planes meet at a common point, which is more likely the more equations are introduced into the system. Examples: Parallel lines, trianglular prism from 3 planes.

Gaussische Eliminationsverfahren

Ideal method for solving a $m \times n$ system of equations, easy to implement algorithmically and works for all dimensions.

Pivot - element on the diagonal of a matrix that has a non 0 coefficient

Rang / rank - number of non 0 pivots, ie (number of rows - number of Kompatibilitaetsbedingungen) - the number of linearly independent rows / columns - the number of dimensions of the output of a linear transformation

Kompatibilitaetsbedingungen - Empty rows at the bottom of the matrix (0 coefficients in one of the equations). If their result is not 0 then there are no solutions for the system. If their result is 0 and the number of equations $\leq$ the number of variables, there are infinite solutions.
Intuition: When thinking of the LGS as superposition, each LHS vector has a 0 component in this dimension, meaning that $\forall x \in \mathbb{R}$ scalar in the Lineare Kombination satisfies the system. Viewing the system with insufficient equations as a system of planes, two planes will intersect along an entire line. In 2D, there would just be a single line, which of course has solutions along its entirety.

Any variables not accounted for due to an all 0 row / no pivot in their column are called *free variables* and can take any real value. TODO: Solidify understanding

Tips:
- Never divide / subtract in Gaussian elimination. Either multiply by $\frac{1}{x}$ or -1. Order is half of the work in maths. - *Vasile Gradinaru*
- Switch rows columns carefully **before** carrying out additions.
- When switching rows to get pivots in the correct place, it is usually best to swap a line with zero pivot with the row that has the largest pivot in that place.

U - Upper (Deutsch: R - Rechts) Matrix - Matrix with 0s under the diagonal and any numbers above it
L - Lower Matrix - Matrix with 0s above the diagonal and any numbers below it
Identity Matrix - Matrix with 0s above and below the diagonal, which only contains 1s
Tridiagonal Matrix - Matrix with 3 diagonals, and otherwise 0s everywhere

Homogene LGS - $\boldsymbol{Ax} = 0$ hat eine triviale Loesung $\boldsymbol{x} = 0$, unless it has free variables.

Square Matrices ($m \times n$):
Regular Matrix, Rank = n, has exactly one solution for arbitrary RHS and only the trivial solution when homogenous
Singular Matrix (Single / peculiar), Rank < n, has infinite / no solutions and has infinite non trivial solutions when homogenous

$m > n$ - An overdetermined LGS only has solutions for specific RHS values (if the rows are not linearly dependent)

Transposed Matrix
For a matrix with notation:

$$i := 1, ..., m \ \text{Zeilen}$$
$$j := 1, ..., n \ \text{Spalten}$$
$$\boldsymbol{A} = \begin{bmatrix} a_{ij} \end{bmatrix}$$
$$\text{Example: } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

The transposed matrix A^T is:

$$\boldsymbol{A^T} = \begin{bmatrix} a_{ji} \end{bmatrix}$$

$$\text{Example: } \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

This can be thought of as pinning the first elements of each row and letting the rest of the row swing down vertically.

Hermitian matrix - $\boldsymbol{A} \in \mathbb{C}^{m \times n}$, $\boldsymbol{A^H}$ - The same procedure however each element becomes its complex conjugate $\bar{a}$.

Vectors may be treated like $\mathbb{R}^{n \times 1}, \mathbb{C}^{n \times 1}$ matrices and transposed in the same manner.

Matrix addition / scalar multiplication is carried out in the same way as vectors.

$$(\boldsymbol{A} + \boldsymbol{B})^T = \boldsymbol{A^T} + \boldsymbol{B^T}$$

$$(\boldsymbol{A} + \boldsymbol{B})^H = \boldsymbol{A^H} + \boldsymbol{B^H}$$

Matrix Symmetry

Symmetrical - $\boldsymbol{A^T} = \boldsymbol{A}$
Antisymmetrical - $\boldsymbol{A^T} = -\boldsymbol{A}$
Hermetian Symmetry - $\boldsymbol{A^H} = \boldsymbol{A}$

Matrix Multiplication

Can be thought as the combination of transformations of space.

Two matrices may only be multiplied if they have the same inner dimensions:

$$\boldsymbol{A_{X \times Y}} \times \boldsymbol{B_{Y \times Z}} = \boldsymbol{C_{X \times Z}}$$

Several LGS with the same LHS can be solved simultaneously with matrix multiplication:

$$\boldsymbol{X} = \left[\overrightarrow{X_1}, ..., \overrightarrow{X_n} \right], \boldsymbol{B} = \left[\overrightarrow{B_1}, ..., \overrightarrow{B_n} \right]$$

$$\boldsymbol{A}^{-1}\boldsymbol{X} = \boldsymbol{B}$$

$\text{Rank}(\boldsymbol{AX}) = \min(\text{Rank}(\boldsymbol{A}), \text{Rank}(\boldsymbol{X}))$

Matrix multiplication is usually not commutative, however always associative and distributive.

Inverse

The inverse of a matrix $\boldsymbol{A}$ is denoted as $\boldsymbol{A}^{-1}$, which reverses the transformation of space represented by matrix $\boldsymbol{A}$. Therefore $\boldsymbol{AA}^{-1} = \boldsymbol{I}$.

The inverse can be used to solve a LGS for arbitrary RHS vectors.

Regulaer, invertierbar und voller Rang sind synonyme dafuer, dass eine Matrix einen Inverse hat. Therefore here are some equivalent conditions which show that a matrix $\boldsymbol{A}$ is regular:

- $\boldsymbol{A}$ is invertierbar
- $\text{Rang}(\boldsymbol{A}) = n$
- $\boldsymbol{Ax} = \boldsymbol{b}$ is solvable for any $\boldsymbol{b}$
- $\boldsymbol{Ax} = \boldsymbol{0}$ only has the trivial solution $x = 0$

Identities:

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$
$$(\boldsymbol{A^T})^{-1} = (\boldsymbol{A}^{-1})^T$$

Eliminationsmatrix (aka Protokolmatrix)

Matrix used for tracking the process of Gaussian elimination. The LHS / RHS multiplied by the elimination matrix results in the current state of the elimination!

It starts as the identity matrix, then the scalar by which another row was multiplied by before adding is written in the position of the currently eliminated variable of the row it was added to. **Caution**: when swapping rows, do NOT forget adjusting the Elimination Matrix accordingly, by simply swapping all non diagonal values in the rows (this is done in a mathematical manner with P matrices later).

Properties of elimination matrices:
- The inverse of the elimination matrix is itself, but non diagonal values become negative. This makes sense intuitively, as $\boldsymbol{EE}^{-1} = \boldsymbol{I}$ so for $\boldsymbol{E}_{ij} + \boldsymbol{E}_{ij}^{-1} = \boldsymbol{0}$ they must have opposite polarities.
- Two lower elimination matrices (with no overlapping elements!) multiplied together is the identity matrix with the combination of both lower elements. This means we can chain steps of Gaussian elimination together nicely. TODO: The conditions for this aren't mentioned in the script, did I understand this correctly? Is it even possible for an elimination matrix that doesn't satisfy these conditions to arise from Gaussian elimination?

Permutation Matrix

Matrix used to track the permutation of rows in LU-Zerlegung. This is simply the identity matrix with the corresponding rows swapped.

Properties of permutation matrices:

$$\boldsymbol{P}_{13} := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{P}^{-1} = \boldsymbol{P}^T$$

$$\text{Row permutation: } \boldsymbol{P}_{13}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

$$\text{Column permutation: } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\boldsymbol{P}_{13} = \begin{pmatrix} 3 & 2 & 1 \\ 6 & 5 & 4 \\ 9 & 8 & 7 \end{pmatrix}$$

Calculating the Inverse

The inverse can be calculated through Gaussian elimination (full Gaussian elimination, ie. with back substitution already carried out so the LHS matrix is the identity matrix) with a RHS of $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix}$ and then finding which $\boldsymbol{X}$ results in $\boldsymbol{X}^{-1}\boldsymbol{b} =$ our eliminated original matrix (by simply reading the coefficients of each component of $\boldsymbol{b}$).

This can be simplified as the so-called **Gauss-Jordan Elimination**. This can be described as the following transformation through regular Gaussian elimination, where the RHS is the elimination matrix (tracks the elimination).

$$[A \mid I] \rightsquigarrow [I \mid A^{-1}]$$

$$\begin{pmatrix} 2 & -1 & 0 & | & 1 & 0 & 0 \\ -1 & 2 & -1 & | & 0 & 1 & 0 \\ 0 & -1 & 2 & | & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & | & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & | & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & | & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{pmatrix}$$

The elimination matrix can be used with any b to apply the steps of elimination, for example when the first row was multiplied by 2 and added to the 2nd row:

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ 2b_1 + b_2 \\ b_3 \end{pmatrix}$$

Cool Fact: Just before each row was multiplied to make the pivots 1, the pivots of the LHS multiplied together is equal to the determinant, which is how computers calculate it for very large matrices. This can also be related to the fact, that a matrix is only invertable if its determinant is non 0 (therefore there are no empty rows / pivots).

LU Zerlegung

A matrix is decomposed into an upper and lower matrix during Gaussian elimination, such that:

$$A = LU$$
$$PA = LU$$

This is the most computationally demanding step of the elimination algorithm.

When the number of RHS we need to solve for is relatively small and the LGS is extremely large, it is more efficient to carry out LU Zerlegung for each system separately rather than to calculate the inverse through Gauss-Jordan elimination.

Proof of $A = LU$:
For steps $1, 2, 3, ..., n$ of Gaussian elimination, the end result of the LHS is an upper matrix.

$$E_n E_{n-1} E_{n-2} ... E_1 A = U$$

The chain of steps (elimination matrices) can be moved to the RHS by multiplying both sides by their inverses due to $EE^{-1} = I$:

$$A = E_n^{-1} E_{n-1}^{-1} E_{n-2}^{-1} ... E_1^{-1} U$$

The chain of elimination steps can of course be represented as a single lower matrix E, therefore:

$$A = E^{-1}U = LU$$

Where E^{-1} is very easy to find (non diagonal elements simply $\times -1$ as mentioned earlier).

Using a combination of Elimination matrices and row + column permutations (these are needed to preserve the diagonal 1s of the resulting elimination matrix), the entire Gaussian elimination process can be encoded as one L matrix.

This is very powerful as the inverses of E and P matrices are easy to find and apply in reverse to the RHS in order to solve the LGS.

The decomposed system can then be used to solve x in another more efficient way to typical Gauss elimination TODO: Understand from script

Upcoming

Determinant - The factor by which a linear transformation (usually represented as a matrix) changes any area / volume in space. Can only be computed for square matrices.

Non-Zero determinant - No information is lost, there is precisely one transformation which reverses the effects on space (inverse matrix)

Next 3B1B Video - Dot products and duality