

Lineare Algebra

Übungsstunde Notizen:

- <https://n.ethz.ch/~jamatter/>
- <https://www.felixgbreuer.com/linalg>

LGS - Linear Gleichung System - linear system of equations

$\mathcal{A}(x) = Ax$ - the corresponding Abbildung von the matrix A. Matrix multiplication is often referred to as a function

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 $e_1 \dots e_n$ 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, triangular 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 Kompatibilitätsbedingungen)
- the number of linearly independent rows / columns - the number of dimensions of the output of a linear transformation

Kompatibilitätsbedingungen - 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 Linear 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.

“Order is half of the work in maths.” - Vasile Gradinaru

Tips:

- Never divide / subtract in Gaussian elimination. Either multiply by $\frac{1}{x}$ or -1 .
- Switch rows columns carefully **before** carrying out additions.
- **Only** add the row who's pivot is currently being considered! Otherwise it is difficult to capture the operation in the elimination matrix (more on this later).
- 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 - $Ax = 0$ hat eine triviale Lösung $x = 0$, unless it has free variables.

Square Matrices ($m \times n$):

This only applies to square matrices

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) and therefore has no inverse (singular).

Transposed Matrix

For a matrix with notation:

$i := 1, \dots, m$ Zeilen

$j := 1, \dots, n$ Spalten

$$\mathbf{A} = [a_{ij}]$$

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

The transposed matrix \mathbf{A}^T is:

$$\mathbf{A}^T = [a_{ji}]$$

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 - $\mathbf{A} \in \mathbb{C}^{m \times n}$, \mathbf{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.

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

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

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

Matrix Symmetry

Symmetrical - $\mathbf{A}^T = \mathbf{A}$

Antisymmetrical - $\mathbf{A}^T = -\mathbf{A}$

Hermetian Symmetry - $\mathbf{A}^H = \mathbf{A}$

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:

$$\mathbf{A}_{X \times Y} \times \mathbf{B}_{Y \times Z} = \mathbf{C}_{X \times Z}$$

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

$$\mathbf{X} = [\vec{X}_1, \dots, \vec{X}_n], \mathbf{B} = [\vec{B}_1, \dots, \vec{B}_n]$$

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

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

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

Inverse

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

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 A is regular:

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

Identities:

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

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

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.

Important: Keep the elimination matrix lower! This means that for the current column, only the current row with 1 in the diagonal may be added to other rows. If this doesn't work, use a permutation.

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 Permutation 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 $EE^{-1} = I$ so for $E_{ij} + E_{ij}^{-1} = 0$ 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.

Permutation Matrix

Orthogonal 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:

$$P_{13} := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

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

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

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

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 $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix}$ and then finding which \mathbf{X} results in $\mathbf{X}^{-1}\mathbf{b} =$ our eliminated original matrix (by simply reading the coefficients of each component of \mathbf{b}).

This can be simplified as the so-called **Gauss-Jordan Elimination**. This can be described as the following transformation through regular Gaussian elimination. All operations happen on both sides in both matrices, unlike LU decomposition.

$$[\mathbf{A} \mid \mathbf{I}] \rightsquigarrow [\mathbf{I} \mid \mathbf{A}^{-1}]$$
$$\left(\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array} \right)$$

The elimination matrix can be used with any \mathbf{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 Lower Upper (LR Left Right) Zerlegung

A matrix can be decomposed into an upper and lower matrix, such that:

$$\mathbf{A} = \mathbf{LU}$$
$$\mathbf{PA} = \mathbf{LU}$$

This can be used to decouple the factorization phase from the actual solving phase in Gaussian elimination. When the number of RHS we need to solve for is relatively small and the \mathbf{A} is extremely large, it is more efficient to carry out LU Zerlegung and the additional steps to solve each system separately rather than to calculate the inverse through Gauss-Jordan elimination.

Proof of $\mathbf{A} = \mathbf{LU}$:

For steps 1, 2, 3, ..., n of Gaussian elimination, the end result of the LHS is an upper matrix.

$$\mathbf{E}_n \mathbf{E}_{n-1} \mathbf{E}_{n-2} \dots \mathbf{E}_1 \mathbf{A} = \mathbf{U}$$

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

$$\mathbf{A} = \mathbf{E}_n^{-1} \mathbf{E}_{n-1}^{-1} \mathbf{E}_{n-2}^{-1} \dots \mathbf{E}_1^{-1} \mathbf{U}$$

The chain of elimination steps can of course be represented as a single lower matrix \mathbf{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.

Using the LU Decomposition

The decomposed system can then be used to solve for x in the following way:

1. $Ax = b$
2. Decompose into the form $PA = LU$
3. Replacing A as PLU, $Lc = Pb$. c can be solved easily thanks to the form of L
4. Based on the above rearrangement, $c = Ux$, in which x can also be solved easily with backsubstitution

LTD: Other decomposition methods (Cholesky etc)

Vector Operations

Euclidean Norm

$$\begin{aligned} x &\in \mathbb{R} \\ |x| &= \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \\ x &\in \mathbb{C} \\ |x| &= \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} \\ \text{Where } |x_1| &= \sqrt{\text{re}(x_1)^2 + \text{im}(x_1)^2} \end{aligned}$$

Dot / skalar Product

$$\begin{aligned} x, y &\in \mathbb{R}^n \\ \langle x, y \rangle &= x_1y_1 + x_2y_2 + \dots + x_ny_n = x^T y \\ x, y &\in \mathbb{C}^n \\ \langle x, y \rangle &= \overline{x_1}y_1 + \overline{x_2}y_2 + \dots + \overline{x_n}y_n = x^H y \end{aligned}$$

The dot product can be used to calculate the angle between two vectors:

$$\begin{aligned} \langle x, y \rangle &= |x||y| \cos(\widehat{(x, y)}) \\ \widehat{(x, y)} &= \arccos \frac{\langle x, y \rangle}{|x||y|} \end{aligned}$$

$$x \perp y \text{ is Orthogonal} \Leftrightarrow \langle x, y \rangle = 0$$

$\frac{\langle x, y \rangle}{|x||y|} \in [-1, 1]$ is proven by the Cauchy-Schwarz inequality.

Thus it can also be used with a unit vector to calculate the projection of a vector in its direction:

$$|u| = 1$$

$$\langle x, u \rangle = |x| \frac{x_u}{|x|}$$

Cross / Vektor product

The cross product is only defined in \mathbb{R}^3 and results in a vector orthogonal to both inputs, whose direction can be determined with the right hand rule.

$$x, y \in \mathbb{R}^3$$

$$x \times y = \begin{vmatrix} e^x & e^y & e^z \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

It can also be used for determining the angle between two vectors in \mathbb{R}^3 :

$$x, y \in \mathbb{R}^3$$

$$x \times y = |x||y|n \sin \theta$$

(n is the unit vector normal to the plane spanned by x, y)

Orthogonale Matrizen

Orthogonal matrix - A square matrix whose columns are perpendicular to each other (dot product 0) and their Euclidean Norms are 1. They do not change lengths or angles - ie they only rotate / reflect space.

The inverse of a rotation / reflection of space is logically its transposition (consider the rotation of base vectors to different axis):

$$Q^T Q = I : Q \text{ is Orthogonal}$$

$$Q^H Q = I : Q \text{ is Unitär}$$

Sei P Orthogonal

$$QP \text{ und } PQ \text{ sind auch Orthogonal}$$

Orthogonal matrix multiplication is still generally not commutative.

Properties of orthogonal matrices

Let A be an orthogonal matrix:

$$\langle Ax, Ay \rangle = \langle x, y \rangle$$

$$|Ax| = |x|$$

$$\widehat{Ax, Ay} = \widehat{x, y}$$

Rotation Matrix

Matrix that rotates space by α degrees anticlockwise:¹

$$R(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Due to the identity $Q^{-1} = Q^T$ rotation anti-clockwise is:

¹Derivation through trigonometry in script by Dr. V Gradinaru

$$R(\alpha)^T = R(-\alpha) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

For a rotation in a certain plane of \mathbb{R}^n , simply keep all columns and rows unaffected by the rotation as the identity matrix. For example a anti-clockwise rotation in the $x \times y$ plane of \mathbb{R}^3 :

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$x \times z$ plane:

$$\begin{pmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{pmatrix}$$

and so on in higher dimensions...

Givens Rotation $G(\varphi) = R(-\varphi)$ - Simply a rotation by the angle φ in the anti-clockwise direction

Rotations which are not confined to a plane can be achieved through a series of plane rotations multiplied together, still resulting in an orthogonal matrix.

Reflection (Householder) Matrix

The orthogonal Householder matrix Q represents the reflection of space (a vector x which does not lie on the plane) over an arbitrary plane with normal **unit** vector u :

$$|u| = 1$$

$$Q = I - 2uu^T$$

“Fuer die Computer sind alle Zahlen schön” - Vasile Gradinaru

QR Zerlegung (QU Zerlegung)

QR Decomposition is a different approach to solving LGSs, where the matrix Q is orthogonal.

Likewise to LU decomposition, it can be used to solve LGSs:

$$Ax = b \Leftrightarrow QRx = b \Leftrightarrow Rx = Q^T b$$

Advantage:

- Reduced rounding errors due to the way computers represent floating point numbers (fractions)

Disadvantage:

- 3 times as inefficient as LU-Zerlegung

The goal is to find a series of orthogonal transformations, which transform A into an upper matrix when multiplying it. Since orthogonal matrices multiplied together result in an orthogonal matrix, they can easily be combined into one and used to solve the LGS as described above.

Givens rotations

Considering a vector, for example $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. We can transform one of its components into 0 by rotating it onto an axis, resulting in $\begin{pmatrix} r \\ 0 \end{pmatrix}$ the following applies:

$$r = \sqrt{x_1^2 + x_2^2}$$

$$\cos(\theta) = \frac{x_1}{r}$$

$$\sin(\theta) = \frac{x_2}{r}$$

This works in \mathbb{R}^n dimensions. TODO: Complex as well?

Using these formulas with a Givens (clockwise) rotation, we can simply replace cos and sin in the matrix with the fractions, targeting the row we don't mind changing and the element we currently want to transform into 0.

By applying a series of Givens rotations, we can carry out the full Gaussian elimination, ending up with an Upper / Right matrix.

Householder reflection

The goal is to find a Householder matrix that reflects a column of A onto the axis of the row we are **not** trying to eliminate. As we know the resulting vector, for example $|x| \cdot e_x$, the unit vector u is simply the normed vector between the original and resulting column of A :

$$x \in \mathbb{R}^n$$

$$v = x - |x| \cdot e_n$$

$$u = \frac{v}{|v|}$$

The Householder matrix can then be found as before:

$$Q = I - 2uu^T$$

When targeting inner columns, the reflection matrix should only be in the bottom right corner of the current Q , to avoid ruining our previous progress (the rest of it stays as I . This means that we ignore the upper row(s).

Which method to choose?

- Givens rotation is great for targeting specific elements to turn into 0, and it needs a series of rotations to reduce several dimensions at once. It is ideal if there are already several 0s in the column.
- Householder reflections have the power to turn all except one element of a column into 0s at once (reflect a vector in \mathbb{R}^3 directly onto the x-axis for example)

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