

Linear algebra: solving systems of linear equations

Statistical Natural Language Processing 1

Çağrı Çöltekin

University of Tübingen
Seminar für Sprachwissenschaft

Winter Semester 2024/2025

© Çöltekin, 2024. 2024-11-13

Quick recap

So far we reviewed:

- Vectors, matrices
- Operations on vectors and matrices
- Dot product
- Matrix multiplication
- Matrices as operators (linear functions / transformations)
- Linearity and linear combinations

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 1 / 30

Today's lecture

- More concepts from linear algebra
 - Solving systems of linear equations
 - Vector independence, matrix rank, vector spaces, span, basis
 - Matrix inverse

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 2 / 30

Solving systems of linear equations

an example

Solve

$$\begin{aligned} x_1 - x_2 &= -1 \\ 2x_1 - x_2 &= 1 \end{aligned}$$

- From the second equation: $x_2 = 2x_1 - 1$
- Substituting this in the second equation:

$$\begin{aligned} x_1 - (2x_1 - 1) &= -1 \\ x_1 - 2x_1 + 1 &= -1 \\ x_1 &= 2 \end{aligned}$$

- $x_2 = 2x_1 - 1 \Rightarrow x_2 = 3$

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 3 / 30

Solving systems of linear equations

Geometric interpretation (1)



- The solution is the intersection of the lines defined by the equations

© Çöltekin, 2024. University of Tübingen

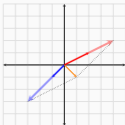
Winter Semester 2024/2025 4 / 30

Solving systems of linear equations

Geometric interpretation (2)

- The solution satisfies the linear combination of the column vectors

$$2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$



© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 5 / 30

Row elimination

$$\begin{aligned} x_1 - x_2 &= -1 \\ 2x_1 - x_2 &= 1 \end{aligned} \iff \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

- We apply a set of *elementary row operations* to the augmented matrix to obtain an *upper triangle matrix*

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

- Elementary row operations are
 - Multiply one of the rows with a non-zero scalar
 - Add (or subtract) a multiple of one row from another
 - Swap two rows

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 6 / 30

Row elimination

an easy example

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

- Add $-2 \times$ row 1 to row 2

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

- This corresponds to:

$$\begin{aligned} x_1 - x_2 &= -1 \\ x_2 &= 3 \end{aligned}$$

where we already see $x_2 = 3$

- *Back-substituting* this in the first equation gives the same answer $x_1 = 2$

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 7 / 30

A (slightly) difficult example

solution is now easy through back-substitution

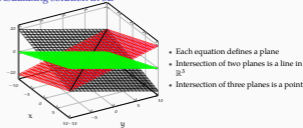
$$\begin{aligned} \begin{bmatrix} 2 & 2 & 4 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 10 \\ 5 \\ 4 \end{bmatrix} \\ \begin{bmatrix} 2 & 2 & 4 & | & 10 \\ 1 & 2 & 1 & | & 5 \\ 1 & 1 & 1 & | & 4 \end{bmatrix} &\rightarrow \begin{bmatrix} 2 & 2 & 4 & | & 10 \\ 0 & 1 & -1 & | & 0 \\ 0 & 0 & -1 & | & -1 \end{bmatrix} \\ 2x_1 + 2x_2 + 4x_3 &= 10 \quad x_3 = 1 \\ x_2 - x_3 &= 0 \quad \rightarrow x_2 = 1 \\ -x_3 &= -1 \quad x_1 = 2 \end{aligned}$$

Can we express the elementary row operations as matrix multiplications?

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 8 / 30

Visualizing solution in 3D



- Each equation defines a plane
- Intersection of two planes is a line in \mathbb{R}^3
- Intersection of three planes is a point

Question: Is this always true?

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 9 / 30

The solution as a linear combination

Our earlier solution

$$\begin{bmatrix} 2 & 2 & 4 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ 4 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

means

$$2 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ 4 \end{bmatrix}$$

Can we solve this equation for any right-hand-side 3-vector?

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 10 / 30

An exercise

Solve,

$$\begin{bmatrix} 4 & 2 & 4 \\ 2 & 2 & 3 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ 4 \end{bmatrix}$$

© Çöltekin, 2024. University of Tübingen

Winter Semester 2024/2025 11 / 30

Singular matrices and matrix rank

- If the elimination results in one or more rows with all zeros, the matrix is said to be *singular*
- This means – effectively – we have fewer equations than unknowns
- If a square matrix is not singular, we can find a unique solution for any right-hand side
- The systems of equations with a singular matrix results in either none or an infinite number of solutions
- The number of columns (or rows) with a pivot is called the **rank** of the matrix
- A non-singular square matrix is said to be full-rank.

A two-dimensional example

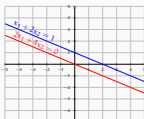
- What is the rank of the following matrix?

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

- Can we solve $Ax = b$
 - for $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$?
 - for $b = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$?

A two-dimensional example

Demonstration of no solution



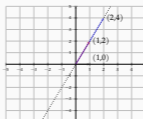
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 2x_1 + x_2 = 1 \\ 4x_1 + 2x_2 = 0 \end{cases}$$

- Lines are parallel to each other: no intersection, no solution

A two-dimensional example

Demonstration of no solution (another view)



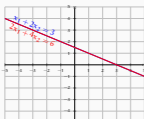
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\Rightarrow x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- All linear combinations of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ bound to be on the dotted line: no linear combination can produce $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

A two-dimensional example

Demonstration of infinite number of solutions



$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 2x_1 + x_2 = 3 \\ 4x_1 + 2x_2 = 6 \end{cases}$$

- Lines are identical: any point on the line is a solution

A two-dimensional example

Demonstration of infinite number of solutions (another view)



$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$$\Rightarrow x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

- There are many (x_1, x_2) combinations that satisfy the equation. An obvious one: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
- More?

Inverse matrix

- If we have a single linear equation with a single unknown: $ax = b$, the solution is

$$x = \frac{1}{a}b \quad \text{or} \quad x = a^{-1}b$$

- We can use an analogous method with systems of linear equations

$$\text{if } Ax = b \text{ then, } x = A^{-1}b$$

- Matrix inverse is only defined for square matrices (not all square matrices are invertible)
- When it exists, $A^{-1}A = AA^{-1} = I$
- If a square matrix is invertible, a version of elimination can be used to find the inverse
 - Create the augmented matrix $[A|I]$
 - Use elementary row operations to obtain $[I|B]$
 - If successful, $B = A^{-1}$

Matrix inversion example/exercise

Invert the following matrix:

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

Properties of matrix inverse

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

LU decomposition

- A square matrix can be factored into two matrices: a lower-triangular matrix L , and an upper-triangular matrix U

$$A = LU$$

- Sometimes a permutation of the original matrix is needed

$$PA = LU$$

- LU decomposition can easily be computed from the results of the row elimination:
 - Elimination gives us U
 - If we keep track of elimination steps, the inverse of the transformations gives L
- LU decomposition useful for many tasks (other than solving systems of linear equations, and finding the inverse of a matrix)

Independence

- A set of vectors v_1, \dots, v_d are *dependent* if any of the vectors can be obtained by a linear combination of others, otherwise they are *independent*
- Alternatively, the column (and row) vectors of a matrix is dependent if $Ax = 0$ has a non-zero solution
- Column (and row) vectors of a square matrix are independent if all columns have a non-zero pivot after row elimination
- Column vectors of a square matrix are independent if and only if row vectors are also independent
- Column vectors of a square matrix are independent if the matrix has full rank
- Column vectors of a square matrix are independent if the matrix has an inverse

Span, basis, and vector spaces

- A set of d independent vectors are said to span a d -dimensional vector (sub-)space. For example,

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

span the whole \mathbb{R}^3

- Any set of vectors that span a vector space forms a *basis* for that vector space
- Any vector in a vector space can be expressed as a linear combination of (any of) the basis for that vector space
- The equation $Ax = b$ has a solution if b is in the vector space spanned by columns of A

Four spaces of a matrix

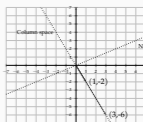
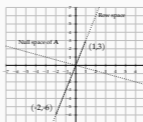
Given a matrix A ,

- **Columns space** of A is the space spanned by the columns of the matrix
- **Row space** of A is the space spanned by the rows of the matrix
- **Null space** of A is the set of vectors x that satisfy $Ax = 0$
 - All vectors in the null space of A are orthogonal to the rows of A
- **Null space** of A^T is the set of vectors x that satisfy $A^T x = 0$, or $x^T A = 0^T$,
 - All vectors in the null space of A^T are orthogonal to the columns of A
- Given an $n \times m$ matrix with rank r
 - Both column and row spaces are r dimensional
 - The dimension of the null space of A is $m - r$
 - The dimension of the null space of A^T is $n - r$

Four spaces of a matrix

$$A = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}$$

A 2x2 example



Systems of equations with rectangular matrices

wide matrices (more columns than rows)

- This means $n \times m$ rectangular matrices with $n < m$,
- Note: the rank of such a matrix is always $\leq n$
- Exercise: solve

$$\begin{bmatrix} 4 & 2 & 4 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \end{bmatrix}$$

- In this case we have
 - no solution if $\text{rank } r < n$ (number of rows)
 - infinitely many solution if $\text{rank } r = n$

Systems of equations with rectangular matrices

tall matrices (more rows than columns)

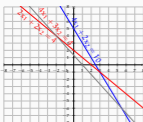
- This means $n \times m$ rectangular matrices with $m < n$,
- Note: the rank of such a matrix is always $\leq m$
- Exercise: solve

$$\begin{bmatrix} 4 & 2 \\ 2 & 2 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ 4 \end{bmatrix}$$

- In this case we have
 - a unique solution if the right-hand side is in the column space of the matrix
 - no solution otherwise
- We will work with this case more often

Visualizing non-solution

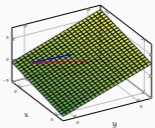
(1) equations as lines in 2-dimensional space



$$\begin{bmatrix} 4 & 2 \\ 2 & 2 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ 4 \end{bmatrix}$$

Visualizing non-solution

(2) column space and the vector b



- The vectors $u = \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}$ and $v = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$ span a 2-dimensional subspace of \mathbb{R}^3
- The vector $w = \begin{bmatrix} 10 \\ 4 \\ 4 \end{bmatrix}$ (scaled to half in the figure) is not on the plane
- We express w as a linear combination of u and v

Summary / next

- Solving sets of linear equations, $Ax = b$, is the focus of linear algebra
- The number of solution depends on the shape and rank of the matrix A
- We also touched on the concepts of
 - independence of sets of vectors
 - vector space
 - basis
 - span
 - matrix rank, column/row/null space

Next:

- Linear regression: trying to solve the unsolvable set of equations

Further reading

Any of the linear algebra references provided earlier.