Linear algebra: solving systems of linear equations

Statistical Natural Language Processing 1

Çağrı Çöltekin

University of Tübingen
Seminar für Sprachwissenschaft

Winter Semester 2025/2026

version: 2f4834d @2025-11-11

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

Today's lecture

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

Linearly (in)dependence

- A set of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is linearly *dependent*, if

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = 0$$

  where $a_1, \ldots, a_n$ not all 0
- If a set of vectors are linearly *dependent* any vector in the set can be written as a linear combination of others
- A set of vectors $\mathbf{v}_1, \mathbf{v}_1, ..., \mathbf{v}_n$ is linearly *independent*, if they are not linearly dependent
- The equation

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n = 0$$

  is true only for $a_i = 0$, for $i = 1, \ldots, n$.

Linearly (in)dependence

Examples

Are the following vectors linearly dependent or independent?

- $\begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}3\\6\end{bmatrix}$ $\quad$ $\begin{bmatrix}-1\\1\end{bmatrix} \begin{bmatrix}2\\2\end{bmatrix}$ $\quad$ $\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}2\\5\end{bmatrix}$
- $\begin{bmatrix}2\\4\\5\end{bmatrix} \begin{bmatrix}3\\1\\1\end{bmatrix} \begin{bmatrix}6\\2\\2\end{bmatrix}$ $\quad$ $\begin{bmatrix}1\\1\\1\end{bmatrix} \begin{bmatrix}2\\1\\1\end{bmatrix} \begin{bmatrix}0\\-1\\-1\end{bmatrix}$ $\quad$ $\begin{bmatrix}1\\1\\1\end{bmatrix} \begin{bmatrix}2\\1\\1\end{bmatrix} \begin{bmatrix}1\\0\\-1\end{bmatrix}$
- If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ are independent, and

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$$

  the set of coefficients $a_1, \ldots, a_n$ are unique
- What is the maximum number of independent n-vectors?

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
- Any set of n independent n-vectors form a basis for $\mathbb{R}^n$

Orthogonal / orthonormal vectors

- A set of orthogonal vectors are independent
- A set of orthogonal unit vectors are called *orthonormal*
- If vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ are orthonormal, then it is easy to find any a_i for

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$$

- Orthogonal / orthonormal vectors have some other useful / interesting properties (we will revisit later)

Solving systems of linear equations

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:

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

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

- We want to make use of the fact that

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

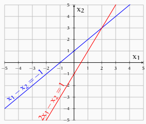
- Rows (equations) define lines in 2D space
- LHS of the equation is a linear combination

$$x_1 \times \begin{bmatrix}1\\2\end{bmatrix} + x_2 \times \begin{bmatrix}-1\\-1\end{bmatrix} = \begin{bmatrix}-1\\1\end{bmatrix}$$

Solving systems of linear equations

Geometric interpretation (1)

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

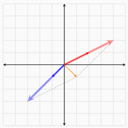


Solving systems of linear equations

Geometric interpretation (2)

- The solution satisfies the linear combination of the column vectors

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



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*

$$\left[\begin{array}{cc|c}1 & -1 & -1\\2 & -1 & 1\end{array}\right]$$

- 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

Row elimination

an easy example

$$\left[\begin{array}{cc|c}1 & -1 & -1\\2 & -1 & 1\end{array}\right]$$

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

$$\left[\begin{array}{cc|c}1 & -1 & -1\\0 & 1 & 3\end{array}\right]$$

- 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 $x_1 = 2$

A (slightly) difficult example

solution is now easy through back-substitution

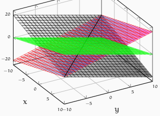
$$\begin{bmatrix} 2 & 2 & 4 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ 4 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 2 & 2 & 4 & 10 \\ 1 & 2 & 1 & 5 \\ 1 & 1 & 1 & 4 \end{array}\right] \quad \left[\begin{array}{ccc|c} 2 & 2 & 4 & 10 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & 1 & 4 \end{array}\right] \quad \left[\begin{array}{ccc|c} 2 & 2 & 4 & 10 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & -1 \end{array}\right]$$

$$\begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & + & 4x_3 & = & 10 \\ & & x_2 & - & x_3 & = & 0 \\ & & & - & x_3 & = & -1 \end{array} \quad \Rightarrow \quad \begin{array}{rcl} x_3 & = & 1 \\ x_2 & = & 1 \\ x_1 & = & 2 \end{array}$$

Can we express the elementary row operations as matrix multiplications?

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?

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?

An exercise (1)

(Try to) solve,

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

An exercise (2)

(Try to) solve,

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

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

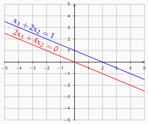
- What is the rank of the following matrix?

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

- Can we solve $\mathbf{A}\mathbf{x} = \mathbf{b}$
  - for $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$?
  - for $\mathbf{b} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$?

A two-dimensional example

Demonstration of no solution



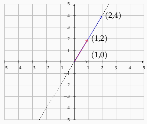
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{array}{rcrcr} 2x_1 & + & x_2 & = & 1 \\ 4x_1 & + & 2x_2 & = & 0 \end{array}$$

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

A two-dimensional example

Demonstration of no solution (another view)



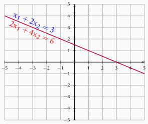
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

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

- 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



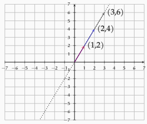
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$$\Rightarrow \begin{array}{rcrcr} x_1 & + & 2x_2 & = & 3 \\ 2x_1 & + & 4x_2 & = & 6 \end{array}$$

- 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

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

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

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

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

LU decomposition

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

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

- Sometimes a permutation of the original matrix is needed

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

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

LU decomposition: example

example: add $-1/2 \times R_1$ to R_3

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \left[\begin{array}{ccc|c} 2 & 2 & 4 & 10 \\ 1 & 2 & 1 & 5 \\ 1 & 1 & 1 & 4 \end{array}\right] = \begin{bmatrix} 2 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix} \times \left[\begin{array}{ccc|c} 2 & 2 & 4 & 10 \\ 1 & 2 & 1 & 5 \\ 1 & 1 & 1 & 4 \end{array}\right] = \left[\begin{array}{ccc|c} 2 & 2 & 4 & 10 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/2 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 2 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 4 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Independence of row and column vectors

- The column (and row) vectors of a matrix is dependent if $\mathbf{A}x = 0$ 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

Four spaces of a matrix

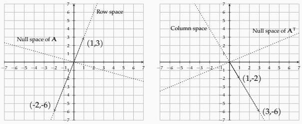
Given a matrix $\mathbf{A}$,

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

Four spaces of a matrix

A 2x2 example

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



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 $\leqslant n$
- 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 rank $r < n$ (number of rows)
  - infinitely many solution if rank $r = n$

Systems of equations with rectangular matrices

tall matrices (more rows than columns)

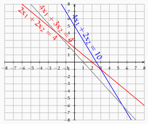
- This means $n \times m$ rectangular matrices with $m < n$,
- Note: the rank of such a matrix is always $\leqslant m$
- Exercise: solve

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

- 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

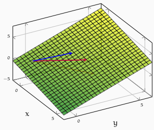
(1) equations as lines in 2-dimensional space



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

Visualizing non-solution

(2) column space and the vector $\mathbf{b}$



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

Summary / next

- Solving sets of linear equations, $\mathbf{Ax} = \mathbf{b}$, is the focus of linear algebra
- The number of solution depends on the shape and rank of the matrix $\mathbf{A}$
- 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.