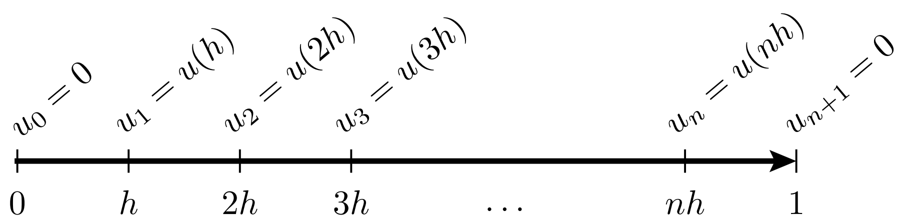
Review

- Goal: solve for $u(x)$ in the **boundary value problem** (BVP)

$$-\frac{\mathrm{d}^2u}{\mathrm{d}x^2} = f(x), \quad 0 \leqslant x \leqslant 1, \qquad u(0) = u(1) = 0.$$

- replace $u(x)$ by its values at equally spaced points in $[0, 1]$



- $-\frac{\mathrm{d}^2u}{\mathrm{d}x^2} \approx -\frac{u(x+h) - 2u(x) + u(x-h)}{h^2}$ at these points (**finite differences**)
- get a linear equation at each point $x = h, 2h, \ldots, nh$; for $n = 5$, $h = \frac{1}{6}$:

$$\underbrace{\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} h^2 f(h) \\ h^2 f(2h) \\ h^2 f(3h) \\ h^2 f(4h) \\ h^2 f(5h) \end{bmatrix}}_{\boldsymbol{b}}$$

- Compute the LU decomposition:

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ -\frac{1}{2} & 1 & & & \\ & -\frac{2}{3} & 1 & & \\ & & -\frac{3}{4} & 1 & \\ & & & -\frac{4}{5} & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & & & \\ & \frac{3}{2} & -1 & & \\ & & \frac{4}{3} & -1 & \\ & & & \frac{5}{4} & -1 \\ & & & & \frac{6}{5} \end{bmatrix}$$

That's how the LU decomposition of **band matrices** always looks like.

Armin Straub
astraub@illinois.edu

LU decomposition vs matrix inverse

In many applications, we don't just solve $A\boldsymbol{x} = \boldsymbol{b}$ for a single $\boldsymbol{b}$, but for many different $\boldsymbol{b}$ (think millions).

Note, for instance, that in our example of "steady-state temperature distribution in a bar" the matrix A is always the same (it only depends on the kind of problem), whereas the vector $\boldsymbol{b}$ models the external heat (and thus changes for each specific instance).

- That's why the LU decomposition saves us from repeating lots of computation in comparison with Gaussian elimination on $[\, A \,|\, \boldsymbol{b} \,]$.
- What about computing A^{-1}?

  We are going to see that this is a bad idea. (It usually is.)

Example 1. When using LU decomposition to solve $A\boldsymbol{x} = \boldsymbol{b}$, we employ forward and backward substitution:

$$A\boldsymbol{x} = \boldsymbol{b} \quad \overset{A=LU}{\Longleftrightarrow} \quad L\boldsymbol{c} = \boldsymbol{b} \quad \text{and} \quad U\boldsymbol{x} = \boldsymbol{c}$$

Here, we have to solve, for each $\boldsymbol{b}$,

$$\begin{bmatrix} 1 & & & & \\ -\frac{1}{2} & 1 & & & \\ & -\frac{2}{3} & 1 & & \\ & & -\frac{3}{4} & 1 & \\ & & & -\frac{4}{5} & 1 \end{bmatrix} \boldsymbol{c} = \boldsymbol{b}, \quad \begin{bmatrix} 2 & -1 & & & \\ & \frac{3}{2} & -1 & & \\ & & \frac{4}{3} & -1 & \\ & & & \frac{5}{4} & -1 \\ & & & & \frac{6}{5} \end{bmatrix} \boldsymbol{x} = \boldsymbol{c}$$

by forward and backward substitution.

How many operations (additions and multiplications) are needed in the $n \times n$ case?

$2(n-1)$ for $L\boldsymbol{c} = \boldsymbol{b}$, and $1 + 2(n-1)$ for $U\boldsymbol{x} = \boldsymbol{c}$.

So, roughly, a total of $4n$ operations.

On the other hand,

$$A^{-1} = \frac{1}{6} \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}.$$

How many operations are needed to compute $A^{-1}\boldsymbol{b}$?

This time, we need roughly $2n^2$ additions and multiplications.

Conclusions

- Large matrices met in applications usually are not random but have some structure (such as band matrices).
- When solving linear equations, we do not (try to) compute A^{-1}.
  - It destroys structure in practical problems.
  - As a result, it can be orders of magnitude slower,
  - and require orders of magnitude more memory.
  - It is also numerically unstable.
  - LU decomposition can be adjusted to not have these drawbacks.

A practice problem

Example 2. Above we computed the LU decomposition for $n=5$. For comparison, here are the details for computing the inverse when $n=3$.

Do it for $n=5$, and appreciate just how much computation has to be done.

Invert $A=\begin{bmatrix} 2 & -1 & \\ -1 & 2 & -1 \\ & -1 & 2 \end{bmatrix}$.

Solution.

$$\left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right] \overset{R2\to R2+\frac{1}{2}R1}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{R3\to R3+\frac{2}{3}R2}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array}\right]$$

$$\overset{\substack{R1\to\frac{1}{2}R1 \\ R2\to\frac{2}{3}R2 \\ R3\to\frac{3}{4}R3}}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array}\right]$$

$$\overset{\substack{R2\to R2+\frac{2}{3}R3 \\ R1\to R1+\frac{1}{2}R2}}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array}\right]$$

Hence, $\begin{bmatrix} 2 & -1 & \\ -1 & 2 & -1 \\ & -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$.

Vector spaces and subspaces

We have already encountered **vectors** in $\mathbb{R}^n$. Now, we discuss the general concept of vectors.

In place of the space $\mathbb{R}^n$, we think of general **vector spaces**.

Definition 3. A **vector space** is a nonempty set V of elements, called **vectors**, which may be added and scaled (multiplied with real numbers).

The two operations of addition and scalar multiplication must satisfy the following *axioms* for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in V, and all scalars c, d.

(a) $\boldsymbol{u} + \boldsymbol{v}$ is in V

(b) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

(c) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

(d) there is a vector (called the **zero vector**) $\mathbf{0}$ in V such that $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$ for all $\boldsymbol{u}$ in V

(e) there is a vector $-\boldsymbol{u}$ such that $\boldsymbol{u} + (-\boldsymbol{u}) = \mathbf{0}$

(f) $c\boldsymbol{u}$ is in V

(g) $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$

(h) $(c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}$

(i) $(cd)\boldsymbol{u} = c(d\boldsymbol{u})$

(j) $1\boldsymbol{u} = \boldsymbol{u}$

tl;dr — A **vector space** is a collection of vectors which can be added and scaled (without leaving the space!); subject to the usual rules you would hope for.

namely: associativity, commutativity, distributivity

Example 4. Convince yourself that $M_{2\times 2} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \text{ in } \mathbb{R} \right\}$ is a vector space.

Solution. In this context, the zero vector is $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Addition is componentwise:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

Scaling is componentwise:

$$r\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}$$

Armin Straub
astraub@illinois.edu

Addition and scaling satisfy the axioms of a vector space because they are defined component-wise and because ordinary addition and multiplication are associative, commutative, distributive and what not.

Important note: we do not use matrix multiplication here!

Note: as a vector space, $M_{2\times 2}$ behaves precisely like $\mathbb{R}^4$; we could translate between the two via

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \longleftrightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

A fancy person would say that these two vector spaces are **isomorphic**.

Example 5. Let $\mathbb{P}_n$ be the set of all polynomials of degree at most $n \geqslant 0$. Is $\mathbb{P}_n$ a vector space?

Solution. Members of $\mathbb{P}_n$ are of the form

$$p(t) = a_0 + a_1 t + \ldots + a_n t^n,$$

where $a_0, a_1, \ldots, a_n$ are in $\mathbb{R}$ and t is a variable.

$\mathbb{P}_n$ is a vector space.

Adding two polynomials:

$$\begin{aligned} & [a_0 + a_1 t + \ldots + a_n t^n] + [b_0 + b_1 t + \ldots + b_n t^n] \\ = \; & [(a_0 + b_0) + (a_1 + b_1) t + \ldots + (a_n + b_n) t^n] \end{aligned}$$

So addition works "component-wise" again.

Scaling a polynomial:

$$\begin{aligned} & r[a_0 + a_1 t + \ldots + a_n t^n] \\ = \; & [(r a_0) + (r a_1) t + \ldots + (r a_n) t^n] \end{aligned}$$

Scaling works "component-wise" as well.

Again: the vector space axioms are satisfied because addition and scaling are defined component-wise.

As in the previous example, we see that $\mathbb{P}_n$ is isomorphic to $\mathbb{R}^{n+1}$:

$$a_0 + a_1 t + \ldots + a_n t^n \quad \longleftrightarrow \quad \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Armin Straub
astraub@illinois.edu

Example 6. Let V be the set of all polynomials of degree exactly 3. Is V a vector space?

Solution. No, because V does not contain the zero polynomial $p(t)=0$.

Every vector space has to have a zero vector; this is an easy necessary (but not sufficient) criterion when thinking about whether a set is a vector space.

More generally, the sum of elements in V might not be in V:

$$[1+4t^2+t^3]+[2-t+t^2-t^3]=[3-t+5t^2]$$

Armin Straub
astraub@illinois.edu