Week 5

LU and LUP decomposition (Section 3.4)

LU decomposition:
Gauss elimination, 3×3, no row exchanges:

$$\underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -c_{32} & 1\end{bmatrix}}_{\substack{\text{subtract } c_{32}\cdot\text{(row 2)}\\ \text{from (row 3)}}} \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -c_{31} & 0 & 1\end{bmatrix}}_{\substack{\text{subtract } c_{31}\cdot\text{(row 1)}\\ \text{from (row 3)}}} \underbrace{\begin{bmatrix}1 & 0 & 0\\ -c_{21} & 1 & 0\\ 0 & 0 & 1\end{bmatrix}}_{\substack{\text{subtract } c_{21}\cdot\text{(row 1)}\\ \text{from (row 2)}}} A = U.$$

Multiplying it out:

$$\overbrace{\begin{bmatrix}1 & 0 & 0\\ -c_{21} & 1 & 0\\ c_{32}c_{21} - c_{31} & -c_{32} & 1\end{bmatrix}}^{L^{-1}\text{, complicated}} A \;=\; U$$

$$\downarrow \text{(take inverse)}$$

$$\underbrace{\begin{bmatrix}1 & 0 & 0\\ c_{21} & 1 & 0\\ c_{31} & c_{32} & 1\end{bmatrix}}_{L\text{, simple}} \qquad\qquad \Rightarrow A = LU$$

Always works: focus on $A \to U$ (Table 3.6 without lines 12 and 22 for $\mathbf{b}$)

Theorem 3.13: Let A be an $m \times m$ matrix on which Gauss elimination succeeds without row exchanges, resulting in an upper triangular matrix U. Let c_{ij} be the multiple of row j that we subtract from row $i > j$ when we eliminate in column j. Then $A = LU$ where

$$L = \begin{bmatrix}1 & & & \\ c_{21} & 1 & & \\ \vdots & & \ddots & \\ c_{m1} & \cdots & c_{m,m-1} & 1\end{bmatrix}.$$

L is lower triangular with 1's on the diagonal.

L is computed "on the side": time still $O(m^3)$.

Proof. Look at a fixed row i. Whenever we change row i, we subtract $c_{ij} \cdot$ (row j) from it, for some previous row j. At this point, row j is "finalized".

$$\begin{array}{r|cccccc|l}
 & u_{11} & \cdots & & & & & \leftarrow \text{finalized (in } U) \\
 & 0 & u_{22} & \cdots & & & & \leftarrow \text{finalized (in } U) \\
 & 0 & 0 & \ddots & & & & \vdots \\
\text{row } j & 0 & 0 & \cdots & \mathbf{u_{jj}} & \cdots & u_{jm} & \leftarrow \text{finalized (in } U) \\
\vdots & & & & & & & \\
\text{row } i & 0 & 0 & \cdots & \star_{ij} & \cdots & \star_{im} & \leftarrow \text{now subtract } c_{ij} \cdot \text{(row } j)
\end{array}$$

What happens to row i?

$$\begin{array}{rccll}
 & & & \text{(row } i) \text{ in } A & \text{initially} \\
- & c_{i1} & \cdot & \text{(row 1) in } U & \text{step 1} \\
- & c_{i2} & \cdot & \text{(row 2) in } U & \text{step 2} \\
\vdots & & & & \\
- & c_{i,i-1} & \cdot & \text{(row } i-1) \text{ in } U & \text{step } i-1 \\
= & & & \text{(row } i) \text{ in } U & \text{finalized.}
\end{array}$$

Move all " $- \cdots$" to the other side: (row i) in A is a linear combination of the first i rows of U. Matrix notation:

$$\text{(row } i) \text{ of } A = \underbrace{\begin{bmatrix} c_{i1} & c_{i2} & \cdots & c_{i,i-1} & 1 & 0 & \cdots & 0 \end{bmatrix}}_{\text{row vector}} U.$$

For all rows of A:

$$A = \underbrace{\begin{bmatrix} 1 & & & \\ c_{21} & 1 & & \\ \vdots & & \ddots & \\ c_{m1} & \cdots & c_{m,m-1} & 1 \end{bmatrix}}_{L} U.$$

$\square$

Solving $A\mathbf{x} = \mathbf{b}$ from $A = LU$:

$$A\mathbf{x} = \mathbf{b}: \quad L \underbrace{U\mathbf{x}}_{\mathbf{y}} = \mathbf{b}.$$

Solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$ (*forward substitution*): $O(m^2)$
Solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$ (back substitution): $O(m^2)$

What if row exchanges are needed?
LU-decomposition may not exist:

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \ell_{11} & 0 \\ \ell_{21} & \ell_{22} \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}}_{U} \text{ has no solution } L, U.$$

LUP decomposition: Official correctness proof of Gauss elimination (Section 3.4.3).
Theorem 3.18: Let A be an $m \times m$ matrix with linearly independent columns, $m \geq 1$. There exist three $m \times m$ matrices P, L, U such that

$$PA = LU,$$

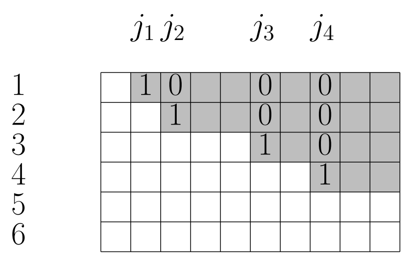
where P is a permutation matrix, L a lower triangular matrix with 1's on the diagonal, and U an upper triangular matrix with nonzero diagonal entries.
Permutation matrix: matrix of linear transformation that reorders the entries of $\mathbf{v}$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_2 \\ v_3 \\ v_1 \end{bmatrix}.$$

L, P are computed "on the side": time still $O(m^3)$.
$A\mathbf{x} = \mathbf{b}$ solved in time $O(m^2)$ for every $\mathbf{b}$.

Gauss-Jordan elimination (Section 3.5)

$A\mathbf{x} = \mathbf{b} \to R_0\mathbf{x} = \mathbf{c}$ with R_0 in row echelon form; works for *every* system!



$\text{REF}(2,3,6,8), r = 4$

Definition 3.19: Let $R = [r_{ij}]_{i=1,j=1}^{m\ \ n}$ be an $m \times n$ matrix. R is in *row echelon form* (REF) if the following holds: There exist $r \leq m$ column indices $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ such that:

(i) For $i = 1, 2, \ldots, r$, we have $r_{ij_i} = 1$ (1's in gray).

(ii) For all i, j, we have $r_{ij} = 0$ whenever $i > r$ (completely white rows) or $j < j_i$ (partially white rows) or $j = j_k$ (0's in gray) for some $k > i$.

If $r = m$, R is in *reduced row echelon form* (RREF) (no completely white rows).
Precise description: $\text{REF}(j_1, j_2, \ldots, j_r)$ or $\text{RREF}(j_1, j_2, \ldots, j_m)$.
Columns $j_1, j_2, \ldots, j_r$: the first r standard unit vectors
I $(m \times m)$: in $\text{RREF}(1, 2, \ldots, m)$
0 $(m \times m)$: in $\text{RREF}()$ $(r = 0)$

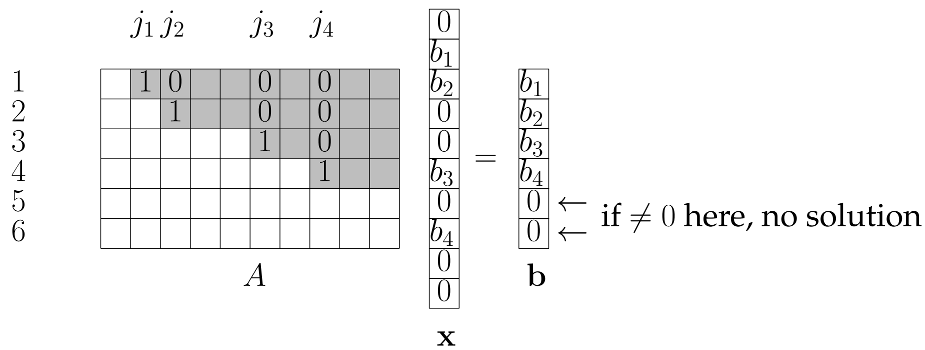
Observation 3.20: A matrix R in $\text{REF}(j_1, j_2, \ldots, j_r)$ has rank r.

Proof. Columns $j_1, j_2, \ldots, j_r$ are the independent ones. □

Direct solution: if A in $\mathrm{REF}(j_1, j_2, \ldots, j_r)$ (rows $i > r$ are zero)

If $b_i \neq 0$ for some $i > m$: no solution! Otherwise:

$$x_j = \begin{cases} b_i, & \text{if } j = j_i \\ 0, & \text{otherwise.} \end{cases} \qquad (\textit{canonical} \text{ solution})$$

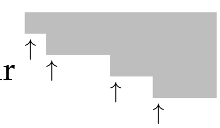


Elimination: if A is not in REF

- $A\mathbf{x} = \mathbf{b} \to R_0\mathbf{x} = \mathbf{c}$ (same solutions, R_0 in REF) focus on $A \to R_0$
- For $R_0\mathbf{x} = \mathbf{c}$, apply direct solution

Like Gauss, except…

…turn pivots into 1: r counts "downward steps" so far



$$A = \begin{bmatrix} \mathbf{2} & 4 & 2 & 2 & -2 \\ 6 & 12 & 6 & 7 & 1 \\ 4 & 8 & 2 & 2 & 6 \end{bmatrix} \quad (r = 0)$$

divide (row 1) by 2:

$$\downarrow \quad \begin{bmatrix} \mathbf{1} & 2 & 1 & 1 & -1 \\ 6 & 12 & 6 & 7 & 1 \\ 4 & 8 & 2 & 2 & 6 \end{bmatrix}$$

subtract $6\cdot$(row 1) from (row 2):

$$\downarrow \quad \begin{bmatrix} \mathbf{1} & 2 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 7 \\ 4 & 8 & 2 & 2 & 6 \end{bmatrix}$$

subtract $4\cdot$(row 1) from (row 3):

$$\downarrow \quad \begin{bmatrix} \mathbf{1} & 2 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & -2 & -2 & 10 \end{bmatrix}$$

downward step made, next column!

$$\downarrow \quad \begin{bmatrix} 1 & 2 & 1 & 1 & -1 \\ 0 & \mathbf{0} & 0 & 1 & 7 \\ 0 & 0 & -2 & -2 & 10 \end{bmatrix} \quad (r = 1)$$

…embrace ugly case: no downward step, next column!

$$\begin{bmatrix}1&2&1&1&-1\\0&0&\mathbf{0}&1&7\\0&0&-2&-2&10\end{bmatrix}\quad(r=1)$$

exchange (row 2) and (row 3): $\downarrow$

$$\begin{bmatrix}1&2&1&1&-1\\0&0&\mathbf{-2}&-2&10\\0&0&0&1&7\end{bmatrix}$$

divide (row 2) by -2: $\downarrow$

$$\begin{bmatrix}1&2&1&1&-1\\0&0&\mathbf{1}&1&-5\\0&0&0&1&7\end{bmatrix}$$

…also eliminate *above* the pivot:

$$\begin{bmatrix}1&2&1&1&-1\\0&0&\mathbf{1}&1&-5\\0&0&0&1&7\end{bmatrix}$$

subtract $1\cdot$(row 2) from (row 1): $\downarrow$

$$\begin{bmatrix}1&2&0&0&4\\0&0&\mathbf{1}&1&-5\\0&0&0&1&7\end{bmatrix}$$

downward step made, next column! $\downarrow$

$$\begin{bmatrix}1&2&0&0&4\\0&0&1&1&-5\\0&0&0&\mathbf{1}&7\end{bmatrix}\quad(r=2)$$

subtract $1\cdot$(row 3) from (row 2): $\downarrow$

$$\begin{bmatrix}1&2&0&0&4\\0&0&1&0&-12\\0&0&0&\mathbf{1}&7\end{bmatrix}$$

m downward steps made, done! $\downarrow$

$$R_0=\begin{bmatrix}1&2&0&0&4\\0&0&1&0&-12\\0&0&0&1&7\end{bmatrix}\quad(r=3)$$

Theorem 3.21: Let A be an $m \times n$ matrix. There exists an invertible $m \times m$ matrix M such that $R_0 = MA$ is in REF.

M: product of (invertible) row operation matrices:

- row exchanges
- row divisions
- row subtractions (below and above the pivot)

Solving $A\mathbf{x} = \mathbf{b}$:

- $A \to R_0 = MA, \mathbf{b} \to \mathbf{c} = M\mathbf{b}$ (like in Gauss, apply row operations also to $\mathbf{b}$)
- $A\mathbf{x} = \mathbf{b}$ and $R_0\mathbf{x} = \mathbf{c}$ have the same solutions (M is "undoable", proof of Lemma 3.3 applies).
- Use direct solution on $R_0\mathbf{x} = \mathbf{c}$.

Lemma 3.22: Let A be an $m \times n$ matrix, M an invertible $m \times m$ matrix, and $R_0 = MA$ in $\mathrm{REF}(j_1, j_2, \ldots, j_r)$. Then A has independent columns $j_1, j_2, \ldots, j_r$.

Proof.
Column j of $\begin{matrix} A \\ R_0 \end{matrix}$ is dependent $\Leftrightarrow$ there is $\mathbf{x}$ in $\mathbb{R}^n$: $\underbrace{\begin{matrix} A\mathbf{x} = \mathbf{0} \\ R_0\mathbf{x} = \mathbf{0} \end{matrix}, x_j = -1, x_k = 0 \text{ for } k > j}_{\text{column } j \text{ is linear combination of previous ones}}$

$\mathbf{x}$ works for $A \Leftrightarrow \mathbf{x}$ works for R_0, since $A\mathbf{x} = \mathbf{0}$ and $MA\mathbf{x} = \mathbf{0}$ have the same solutions (proof of Lemma 3.3 with $\mathbf{b} = \mathbf{0}$). A and R_0 have the same (in)dependent columns.

R_0 has independent columns $j_1, j_2, \ldots, j_r$ (Observation 3.20). Therefore, A has the same. □

If A is $m \times m$, invertible:
all columns are independent $\Rightarrow R_0 = MA$ in $\mathrm{RREF}(1, 2, \ldots, m) \Rightarrow R_0 = I \Rightarrow M = A^{-1}$.

Computing the CR decomposition:
Recall Theorem 2.23:

$$A = \underbrace{C}_{m\times r}\underbrace{R}_{r\times n}.$$

C submatrix of independent columns; R how to combine them to get all columns.

Theorem 3.24: Let A be an $m \times n$ matrix, $A = CR$ (according to Theorem 2.23), $A \to R_0 = MA$ in $\mathrm{REF}(j_1, j_2, \ldots, j_r)$. Then

- $R =$ the first r rows of R_0 (the nonzero rows of R_0).
- $C =$ columns $j_1, j_2, \ldots, j_r$ of A (the independent columns of A)

Proof. $R_0 = M\underbrace{CR}_{A}$.

- C has columns $j_1, j_2, \ldots, j_r$ of A (the independent ones by Lemma 3.22).
- MC has columns $j_1, j_2, \ldots, j_r$ of $R_0 = MA$: the unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$.

$$R_0 = MCR = \underbrace{\left[\begin{array}{c} \underbrace{I}_{r\times r} \\ \hline \underbrace{0}_{(m-r)\times r} \end{array}\right]}_{MC} R = \underbrace{\left[\begin{array}{c} \underbrace{R}_{r\times n} \\ \hline \underbrace{0}_{(m-r)\times n} \end{array}\right]}_{R_0}.$$

□

Verify this on

$$\underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 4 \\ 3 & 6 & 2 & 5 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}}_{R}$$

from Section 2.2.3 by doing Gauss-Jordan on A!