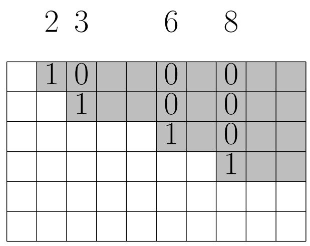
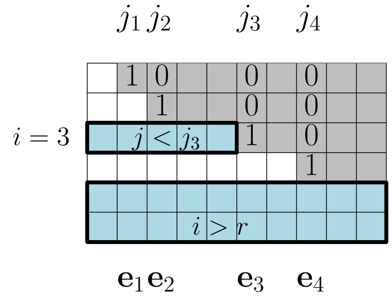
Week 6

Gauss-Jordan elimination (Section 3.3)

$A\mathbf{x} = \mathbf{b} \to R\mathbf{x} = \mathbf{c}$ with R in reduced row echelon form (RREF); works for *every* system!



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



$r_{ij} = 0$ whenever

$i \le r$ and $j < j_i$

or

$i > r$

In general: $\text{RREF}(j_1, j_2, \ldots, j_r)$: $j_1, j_2, \ldots, j_r$ the "downward step" columns (Definition 3.13)

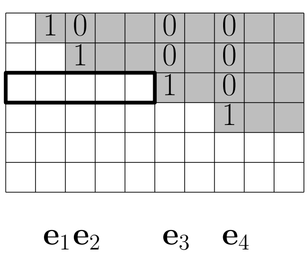
I ($m \times m$): in $\text{RREF}(1,2,\ldots,m)$

0 ($m \times m$): in $\text{RREF}()$ ($r = 0$)

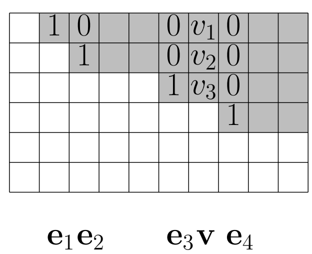
Lemma 3.14: R in $\text{RREF}(j_1, j_2, \ldots, j_r)$ has independent columns $j_1, j_2, \ldots, j_r$ and rank r.

"Proof" by picture:

independent ↓ : only 0's left of 1

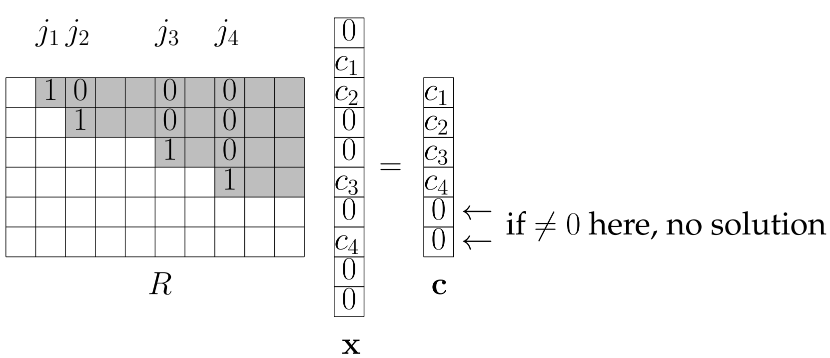


dependent ↓



$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ 0 \\ 0 \\ 0 \end{pmatrix} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3$$

Direct solution of $R\mathbf{x} = \mathbf{c}$ with R in $\text{RREF}(j_1, j_2, \ldots, j_r)$:



$$x_j = \begin{cases} c_i, & \text{if } j = j_i \\ 0, & \text{otherwise.} \end{cases}$$

canonical solution

Elimination: if A is not in RREF

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

Like Gauss, except...turn pivots into 1, also eliminate *above* the pivots: Column 1

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

divide (row 1) by 2: $\downarrow$

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

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

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

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

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

Pivot **0** and no row exchange possible. Column 2 $\to$ Column 3

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

$$\downarrow$$

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

Case of failure in Gauss elimination. Here: case that saves us some work!

Column 3

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

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: $\downarrow$

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

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

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

Column 4

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

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

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

What about the right-hand side $\mathbf{b}$?

Very useful: Version with m right-hand sides as input (columns of $m \times m$ matrix B) and the m transformed right hand-sides as output (columns of C).

Algorithm 6:

1: **function** GAUSS-JORDAN ELIMINATION(A, B) ▷ $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times m}$
2: $R \leftarrow A, C \leftarrow B, r \leftarrow 0$ ▷ r: number of downward steps so far
3: **for** $j = 1, 2, \ldots, n$ **do** ▷ eliminate in column j
4: **if** $r = m$ **then** ▷ no further downward steps possible, done!
5: **break** ▷ …loop execution and go to line 21
6: **end if**
7: $s \leftarrow r + 1$ ▷ row of (potential) next downward step
8: **if** $r_{sj} = 0$ **then** ▷ zero pivot
9: **if** there is some $k > s$ such that $r_{kj} \neq 0$ **then**
10: exchange (row s) and (row k) (in both R and C) ▷ row operation

```
11:         else                                          ▷ no downward step in column j
12:             continue                                  ▷ ... in line 3 with the next column
13:         end if
14:     end if                                            ▷ now r_sj ≠ 0
15:     divide (row s) by r_sj (in both R and C)          ▷ row operation
16:     for i = 1, 2, ..., s − 1 and i = s + 1, s + 2, ..., m do   ▷ make r_ij = 0
17:         subtract r_ij· (row s) from (row i) (in both R and C)  ▷ row operation
18:     end for                                           ▷ now, the j-th column of R equals e_s
19:     r ← r + 1, j_r ← j                                ▷ next downward step was made in column j
20:   end for
21:   return (R, j_1, j_2, ..., j_r, C)                   ▷ R is in RREF(j_1, j_2, ..., j_r)
22: end function
```

Runtime for $(A, B) \to (R, C)$:

- $O(m)$ row operations in each of the r downward steps ($r \leq m$)
- $O(m+n)$ basic operations per row operation (C: m columns, A: n columns)
- Time $O(mr(m+n))$ in total (also covers the other operations)

For $m = r = n$ (Gauss success scenario): $O(m^3)$.

Theorem 3.17: Let A be an $m \times n$ matrix, and let $(R, j_1, j_2, \ldots, j_r, M)$ be the output of Algorithm 6 with input (A, I). Then M is invertible, $R = MA$, and R is in $\mathrm{RREF}(j_1, j_2, \ldots, j_r)$.

Proof. $\underbrace{C}_{M} = \underbrace{M_\ell M_{\ell-1} \ldots M_1}_{\text{row operation matrices}} \underbrace{B}_{I} \quad \Rightarrow \quad R = \underbrace{M_\ell M_{\ell-1} \ldots M_1}_{M} A.$
Since all M_i are invertible (row operations are undoable), their product M is also invertible (Lemma 2.59). □

R is the "RREF standard form" of A and also gives us the CR decomposition:

Theorem 3.18: Let A be an $m \times n$ matrix. There is a unique $m \times n$ matrix R (the one resulting from Gauss-Jordan elimination; Theorem 3.17), with the following two properties.

(i) $R = MA$ for some invertible $m \times m$ matrix M.

(ii) R is in RREF.

More precisely, R is in $\mathrm{RREF}(j_1, j_2, \ldots, j_r)$, where $j_1, j_2, \ldots, j_r$ are the indices of the independent columns in A, and

$$R = \left[\begin{array}{c} \underbrace{R'}_{r \times n} \\ \hline \underbrace{0}_{(m-r) \times n} \end{array}\right],$$

with R' the unique matrix such that $A = CR'$ in Theorem 2.46 (CR decomposition).

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 2.3.5 by doing Gauss-Jordan on A:

$$\begin{array}{rr} & A = \begin{bmatrix}1 & 2 & 0 & 3\\ 2 & 4 & 1 & 4\\ 3 & 6 & 2 & 5\end{bmatrix}\\ \text{elimination in column 1:} & \downarrow\\ & \begin{bmatrix}1 & 2 & 0 & 3\\ 0 & 0 & 1 & -2\\ 0 & 0 & 2 & -4\end{bmatrix}\\ \text{elimination in column 3:} & \downarrow\\ & R = \begin{bmatrix}1 & 2 & 0 & 3\\ 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 0\end{bmatrix}\end{array}$$

Computing the inverse of A: Run Gauss-Jordan on $(A, I) \to (R, M)$.
Theorem 3.19: A invertible $\Leftrightarrow R = I$ (and then $M = A^{-1}$).

Solving $A\mathbf{x} = \mathbf{b}$ (possibly for many b): Run Gauss-Jordan on $(A, I) \to (R, M)$.
Theorem 3.20: Set $\mathbf{c} = M\mathbf{b}$ and solve $R\mathbf{x} = \mathbf{c}$ ($MA\mathbf{x} = M\mathbf{b}$) by direct solution.

$O(m^2)$ time per **b**, after $O(mr(m+n))$ time for Gauss-Jordan.

Vector spaces (Section 4.1)

`https://ti.inf.ethz.ch/ew/courses/LA25/slides/vectors_handout.pdf`

Bases and dimension (Section 4.2)

Basis of a vector space V: linearly independent vectors that span V.

Need linear combinations / span in vector spaces.

Previously, we used *sequences* of vectors; now, we also work with *sets* (possibly infinite: polynomials).

Definition 4.15 Let V be a vector space, $G \subseteq V$ a (possibly infinite) subset of vectors. A *linear combination* of G is a sum of the form

$$\sum_{j=1}^{n} \lambda_j \mathbf{v}_j,$$

where $F = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a finite subset of G and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$.
Lemma 4.16: Let V be a vector space. Every linear combination of V is again in V.

Proof is easy (linear combination is "natural behavior", combining "add up" and "scale"). But needs *finite* linear combinations:

Consider $\mathbb{R}[x]$ (polynomials), and the infinite "linear combination"

$$\sum_{j=0}^{\infty} x^j \left(= \frac{1}{1-x}\right).$$

of the *unit monomials* $1, x, x^2, \ldots$. This is not a polynomial.

Definition 4.17: Let V be a vector space, $G \subseteq V$ (possibly infinite).

G is called *linearly dependent* if there is an element $\mathbf{v} \in G$ such that $\mathbf{v}$ is a linear combination of $G \setminus \{\mathbf{v}\}$. Otherwise, G is called *linearly independent*.

The *span* of G, $\mathbf{Span}(G)$, is the set of all linear combinations of G.

Bases

Definition 4.18: Let V be a vector space. A subset $B \subseteq V$ is called a *basis* of V if B is linearly independent and $\mathbf{Span}(B) = V$.

Examples: (For linear independence, use *private nonzero* argument!)

vector space V	basis B
$\mathbb{R}^m$	$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$
$\mathbf{C}(A)$ (subspace of $\mathbb{R}^m$)	independent columns of A
2×2 symmetric matrices (subspace of $R^{2\times 2}$)	$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$
$\mathbb{R}[x]$ (polynomials)	$\{x^i : i = 0, 1, \ldots\}$ (infinite set)
$\{\mathbf{0}\}$ (smallest vector space)	$\emptyset$ (empty set)

There can be many bases:
Observation 4.20: Every set $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ of m linearly independent vectors is a basis of $\mathbb{R}^m$.

Proof. B is linearly independent, and $\mathbf{Span}(B) = \mathbb{R}^m$ is true by Lemma 1.28. □

Does every vector space have a basis? Yes, but we only treat the *finitely generated* case.

Definition 4.21: A vector space V is called *finitely generated* if there exists a finite subset $G \subseteq V$ with $\mathbf{Span}(G) = V$.

$\mathbb{R}^m$ is finitely generated (by $G = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$) but $\mathbb{R}[x]$ is not.

Theorem 4.22: Let V be a finitely generated vector space, and $G \subseteq V$ a finite subset with $\mathbf{Span}(G) = V$. Then V has a basis $B \subseteq G$.

(Algorithmic) Proof. Start with G.

1. If G is linearly independent, G is a basis.

2. Otherwise, some $v \in G$ is a linear combination of $G \setminus \{v\}$ (Definition 4.17) and $\mathbf{Span}(G \setminus \{\mathbf{v}\}) = \mathbf{Span}(G) = V$ (Corollary 1.27).

3. Replace G with $G \setminus \{\mathbf{v}\}$ and goto 1.

G is finite and gets smaller in every round, so this must stop with a basis. □

The Steinitz exchange lemma

Lemma 4.23: Let V be a finitely generated vector space, $F \subseteq V$ a finite set of linearly independent vectors, and $G \subseteq V$ a finite set of vectors with $\mathbf{Span}(G) = V$. Then the following two statements hold.

(i) $|F| \leq |G|$.

(ii) There exists a subset $E \subseteq G$ of size $|G| - |F|$ such that $\mathbf{Span}(F \cup E) = V$.

(ii) means: can enlarge F by some elements from G such that the result has at most the size of G and also spans V.

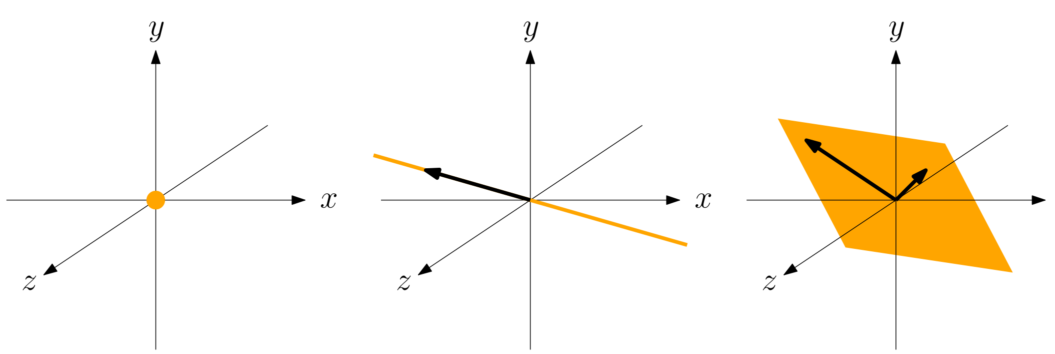
This is the luxury version of Lemma 1.28. Proof for both is (almost) the same:
`https://scratch.mit.edu/projects/1226855322/`.

Theorem 4.24 Let V be a finitely generated vector space and let $B, B' \subseteq V$ be two bases of V. Then $|B| = |B'|$.

Proof. B and B' are linearly independent, and $\mathbf{Span}(B) = \mathbf{Span}(B') = V$ (Definition 4.18). Steinitz exchange Lemma 4.23 (i) with $F = B, G = B'$ gives $|B| \leq |B'|$; with $F = B', G = B$, we get $|B'| \leq |B|$. □

Dimension

Definition 4.25: Let V be a finitely generated vector space. Then the *dimension* of V, $\dim(V)$, is the size of an arbitrary basis B of V.



dimension 0 dimension 1 dimension 2

Now we can finally say: $\mathbb{R}^m$ has dimension m. :-)