# MA251 Linear algebra II :: Lecture notes

Lecturer: Christian Boehning

November 2, 2023

## Contents

Topics

- Spectral theory of endomorphisms
- Bilinear forms
- Excursion into linear algebra over $\mathbb{Z}$

Literature

- Lecture notes on Moodle
- Peter Lax. Linear algebra and its applications. Wiley

# 1 Reminders from MA106

## 1.1 Field $K$

is a set with 2 maps:

- "+" $K \times K \to K$
- "$\cdot$" $K \times K \to K$

satisfying

1. $a + (b + c) = (a + b) + c$ additive associativity
2. $\exists 0 \in K : \forall a \in K, 0 + a = a$ unique $0_k$
3. $\forall a \in K$, $\exists (-a) \in K : a + (-a) = 0$ additive inverse
4. $a + b = b + a$ additive commutativity
5. $(ab)c = a(bc)$ multiplicative associativity
6. $\exists 1 \in K : \forall a \in K, 1 \cdot a = a$ unique $1_k$ $[0_k \neq 1_k]$
7. $\forall a \in K\{0\}$, $\exists a^{-1} : a \cdot a^{-1} = 1$ multiplicative inverse
8. $ab = ba$ multiplicative commutativity
9. $a(b + c) = ab + ac$ distributive law

or that it's an abelian group on + and $\cdot$, which are connected by the distributive law.

**Example 1.1.1.** $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}/p\mathbb{Z}$ $(\mathbb{F}_p)$ where $p$ is prime are fields. $\mathbb{F}_2$ is the smallest field, thus defined:

| $+$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| $\times$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

## 1.2 Vector space $V$ over field $K$

is a set with 2 maps

- "+" $V \times V \to V$
- "$\cdot$" $K \times V \to V$ scalar multiplication

satisfying

1. $(V, +)$ is abelian
2. $\lambda(\mu v) = (\lambda\mu)v$
3. $\forall v \in V$, $1 \cdot v = v$
4. $(\lambda + \mu)v = \lambda v + \mu v$
5. $\lambda(v + w) = \lambda v + \lambda u$ 2 forms of distributive laws

**Example 1.2.1.** $K^n$, $n \in \mathbb{N}$ ($n$-tuples with elements of $K$)

**Definition 1.2.2.** A *basis* of $V$ is a subset $B \subset V$ such that any $v \in V$ is uniquely written as a linear combination of elements of $B$. If $B$ is finite, all basis have the same cardinality, called *dimension* of $V$, $\dim V$.

## 1.3 Linear map $T : V \to W$

is a map such that $\forall \alpha, \beta \in K, v, w \in V$, $T(\alpha v + \beta w) = \alpha Tv + \beta Tw$. [structure preserving]

**Remark.** A lesson given by 20th century mathematics is that maps between objects are sometimes far more important than the objects themselves.

**Example 1.3.1.**
1. $V = W = \mathbb{R}$, $T : x \mapsto x^2$ is not linear, but if $V = W = \mathbb{F}_2$ it would be. [in fact, identity]
2. $V = C([0, 1])$ (continuous functions on $[0, 1]$), $W = \mathbb{R}$. Then $T : \int_0^1$ is linear.
3. $V = C^1(\mathbb{R})$ (continuously differentiable functions on $\mathbb{R}$), $W = C(\mathbb{R})$. Then $T : \frac{\mathrm{d}}{\mathrm{d}x}$ is linear.

The *matrix* of $T : V \to W$ where $V, W$ are $n$ and $m$-dimensional $K$-vector spaces with respect to bases $E, F$, notation $\mathcal{M}(T)_E^F$, is $m \times n$ and its entries $a_{ij}$ satisfy

$$T(e_j) = \sum_{i=1}^{m} a_{ij} f_i,$$

and

$$w_F = \mathcal{M}(T)_E^F \cdot v_E,$$

which actually makes matrix notation make sense.

Matrix multiplication corresponds to composition of linear maps.
With $U_A \xrightarrow{R} V_B \xrightarrow{S} W_C$ we can write

$$\mathcal{M}(S \circ R)_A^C = \mathcal{M}(S)_B^C \cdot (R)_A^B.$$

**Example 1.3.2.** $V = K[x]_{\leq 2}$ (polynomials of deg $\leq 2$ with coefficients in $K$), $W = K[x]_{\leq 1}$, $T : \frac{\mathrm{d}}{\mathrm{d}x}$. Then $\mathcal{M}(T)_E^F$ where $E = (1, x, x^2)$, $F = (x, 2)$ is

$$\begin{pmatrix} 0 & 0 & 2 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

## 1.4 Change of basis

Question: given linear map $T : V \to W$, $\dim V = n$, $\dim W = m$, we have $\mathcal{M}(T)_E^F = A$ and another pair of bases $E', F'$ and $\mathcal{M}(T)_{E'}^{F'} = B$. How are $A$ and $B$ related?
We can see it this way:

$$V_{E'} \xrightarrow{\mathrm{id}} V_E \xrightarrow{T} W_F \xrightarrow{\mathrm{id}} W_{F'},$$

so

$$\mathcal{M}(T)_{E'}^{F'} = \mathcal{M}(\mathrm{id})_F^{F'} \mathcal{M}(T)_E^F \underbrace{\mathcal{M}(\mathrm{id})_{E'}^E}_{\text{writing vectors of } E' \text{ in those of } E} \qquad \text{i.e.} \quad B = Q^{-1}AP.$$

Particularly, given $T : V \to V$ and bases $E$, $E'$, we say $A$ and $B$ are similar if we can write $B = P^{-1}AP$ where $P = \mathcal{M}(\mathrm{id})_{E'}^E$.

## 1.5 Eigen-stuff

**Definition 1.5.1.** *Eigenvalue* is a $\lambda \in K$ such that $\exists v \neq 0 : Tv = \lambda v$. $v$ is then a corresponding *eigenvector*, which defines an invariant subspace.

**Proposition 1.5.2.** If $\lambda_1 \dots \lambda_r$ are pairwise distinct eigenvalues of $T : V \to V$, then corresponding eigenvectors $v_1 \dots v_r$ are linearly independent.

*Proof.* Suppose $v_1 \dots v_r$ are linearly dependent. Then $\exists a_1 v_1 + \dots + a_r v_r = 0$ where not all $a_i = 0$. Choose one that involves minimum number of $v$'s [well-ordering principle] :

$$b_1 v_{i_1} + \dots + b_s v_{i_s} = 0, \quad s \leq r. \qquad (*)$$

Applying $T$ to $*$,

$$\lambda_{i_1} b_1 v_{i_1} + \dots + \lambda_{i_s} b_s v_{i_s} = 0. \qquad (**)$$

while multiplying $*$ by $\lambda_{is}$ and subtracting $**$ give us

$$(\lambda_{i_1} - \lambda_{is}) b v_{i_1} + \dots + (\lambda_{i_{s-1}} - \lambda_{i_s}) b v_{i_{s-1}} = 0.$$

Since $\lambda_i$'s are distinct, the coefficients are non-zero, contradicting that $s$ is minimum. Therefore linear independency is proved. □

**Corollary 1.5.3.** $T : V \to V$ has $\dim V$ distinct eigenvalues $\Leftrightarrow$ it has a basis consisting of eigenvectors $\Leftrightarrow$ has a diagonal matrix (diagonalisable).

# 2 Spectral theory of endomorphisms, Jordan canonical form

- Question 1: Given 2 $n \times n$ matrices $A$, $B$. When are they similar?
- Question 2: Given and $n$-dimensional $K$-vector space $V$ and a linear map $T : V \to V$. Is there a basis $E'$ with respect to which $\mathcal{M}(T)^{E'}_{E'}$ has a particularly *simple* form?

Answer to Q1 are straight forward when $A$, $B$ are diagonalisable: they are similar if and only if their eigenvectors are the same: $\lambda_1 = \mu_1, \lambda_2 = \mu_2, \ldots \lambda_r = \mu_r$ and the dimensions of corresponding eigenspaces are the same: $\dim \mathrm{Ker}\ (A - \lambda_j I_n) = \dim \mathrm{Ker}\ (B - \mu_j I_n)$.

What if they are not diagonalisable? We need stronger condition.

**Definition 2.0.1.** Let $A$ be a $n \times n$ matrix and $\lambda_j$ an eigenvalue of $A$. Then for $i \in \mathbb{N}$, $\mathrm{Ker}\ \left((A - \lambda_j I_n)^i\right)$ is called the *generalised eigenspace* of index $i$, denoted $N_i(A_j, \lambda)$. A nonzero vector in this space is then called a *generalised eigenvector*.

So

- the general answer to Q1 is that generalised eigenvectors are the same and dimensions of corresponding generalised eigenspaces are the same, $\forall j, i \in \mathbb{N}$;
- the general answer to Q2 is that $V$ has $\dim V$ distinct generalised eigenvectors $\Leftrightarrow$ has a basis of generalised eigenvectors of $T$.

## 2.1 Minimal polynomial

We have the characteristic polynomial $c_A(x) = \det(A - xI)$, but it's not able to detect diagonalisability. e.g., given $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, we have $c_A(x) = c_B(x) = (1 - x)^2$, while $A$ is diagonalisable (it is already diagonal) and $B$ is not. But minimal polynomial $\mu_A(x)$ will be; it also classifies nilpotent endomorphisms (a power of which is zero).

**Notation.** Given a polynomial $p \in \mathbb{C}[x]$, we have $p(A)$ where $A \in \mathbb{C}^{n \times n}$ which is a matrix: $\sum_{j=0}^{n} c_j A^j$.

**Definition 2.1.1.** A polynomial is called *monic* if the leading coefficient is 1.

**Definition 2.1.2.** The *minimal polynomial* $\mu_A(x)$ is the unique, nonzero monic polynomial of the smallest degree such that it annihilates the matrix: $\mu_A(A) = 0$.

For this assertion to make sense we need to prove a couple of things.

**Theorem 2.1.3.** Given $A \in K^{n \times n}$, there exists a nonzero polynomial $p$ of degree at most $n^2$ such that $p(A) = 0$.

*Proof.* Note that $\dim K^{n \times n} = n^2$, so the $n + 1$ vectors $I_n, A, A^2, \ldots, A^{n^2}$ is linearly dependent (has a nontrivial linear combination equal to 0), which gives us a polynomial. □

$n^2$ is actually too big, we can have a better bound.

**Theorem 2.1.4** (Cayley–Hamilton)**.** Given $A \in K^{n \times n}$, $c_A(A) = 0$.

**Remark** (before the proof)**.** Suppose we have 2 polynomials $P(x) = \sum_j P_j x^j$, $Q(x) = \sum_j Q_j x^j$ with coefficients $n \times n$ matrices, we can multiply them using the usual definition

$$R(x) = P(x)Q(x) = \sum_l R_l x^l \text{ where } R_l = \sum_{j+k=l} P_j Q_k.$$

If $A \in \mathbb{C}^{n\times n}$ is such that it commutes with all $Q_k$ ($AQ_k = Q_k A$), then we can substitute $A$ in without causing any harm: $R(A) = P(A)Q(A)$.

Recall the adjoint matrix adj $M$ which satisfies $M \cdot \text{adj } M = \text{adj } M \cdot M = \det M \cdot I_n$, with entries adj $M_{ij} = (-1)^{i+j} \det \underbrace{M_{ij}}_{M \text{ with row } j \text{ and column } i \text{ deleted}}$. This makes sense for matrices with entries in any commutative ring, e.g. polynomials.

*Proof.* Let $Q(x) := A - xI_n$ and $P(x) := \text{adj } Q(x)$. Then $P(x)Q(x) = \det(A - xI_n) \cdot I_n = c_A(x)I_n$. Therefore

$$P(A)Q(A) = P(A)(A - AI_n) = 0 = c_A(A)I_n \Rightarrow c_A(A) = 0.$$

□

It's tempting to substitute $A$ for $x$ in $\det(A - xI_n)$ which is of course 0, however $\in K$ instead of $K^{n\times n}$, i.e. a scalar instead of a matrix.

Now recall the definition of minimal polynomial and here's a sanity check for well-definedness.

1. Set of nonzero $p$'s such that $p(A) = 0$ is nonempty by Cayley–Hamilton. Thus by well-ordering principle there is a $p$ of minimal degree.
2. To make it monic we just scale it.
3. Uniqueness: if there were 2 distinct such $p, p'$ both nonzero, monic, of minimal degree $d$, then $p - p'$ is still nonzero but of smaller degree satisfying $p - p'(A) = p(A) - p'(A) = 0$, a contradiction.

**Proposition 2.1.5.** If $q$ is any nonzero polynomial such that $q(A) = 0$ then $\mu_A$ divides $q$.

*Proof.* There is division with remainder (Euclidean algorithm for polynomials): $q = s\mu_A + r$ where $\deg r < \deg p$ or is 0. Since $A$ satisfies $p(A) = 0$ and $\mu_A(A) = 0$, it must be that $r(A) = 0$. Hence $r = 0$ since otherwise it's a contradiction to minimality of $\deg \mu_A$. □

**Proposition 2.1.6.** $\mu_A$ and $c_A$ have the same roots, not counting multiplicities.

*Proof.* We have $c_A(A) = 0$ hence $\mu_A | c_A$. Therefore $\{\text{roots of } \mu_A\} \subset \{\text{roots of } c_A\}$.

Now suppose $\lambda \in \{\text{roots of } c_A\}$, i.e. an eigenvalue of $A$, then $\forall p \in \mathbb{C}[x]$, $p(A)v = p(\lambda)v$. (exercise for assignment) Take $p = \mu_A$, then

$$\mu_A(\lambda)v = \mu_A(A)v = 0, v \neq 0 \Rightarrow \mu_A(\lambda) = 0,$$

i.e. $\{\text{roots of } c_A\} \subset \{\text{roots of } \mu_A\}$. □

**Example 2.1.7.** $\mu_D(x)$ of a diagonal matrix

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}.$$

We know that given any $p \in \mathbb{C}[x]$,

$$p(D) = \begin{pmatrix} p(d_1) & & 0 \\ & \ddots & \\ 0 & & p(d_n) \end{pmatrix}.$$

To make the degree minimal, suppose $\{d_1, \ldots, d_n\} = \{d_{i_1}, \ldots, d_{i_k}\}$ with $d_{i_j}$'s pairwise distinct. Then

$$\mu_D(x) = (x - d_{i_1})(x - d_{i_2}) \cdots (x - d_{i_k}).$$

**Proposition 2.1.8.** $A$ is diagonal (or more generally $A$ is diagonalisable) $\Rightarrow \mu_A$ has no multiple roots.

We'll see later that these two are actually $\Leftrightarrow$. To say "diagonalisable" is fine since $p(SAS^{-1}) = S_\phi(A)S^{-1}$.

**Example 2.1.9.** In the previous example that $I_2$ is diagonalisable and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not, we have that $\mu_{I_2} = x - 1$ and $\mu_B = (x-1)^2$.

Question now is of course how to compute this $\mu_A$.

## 2.2 Methods to calculate $\mu_A$

**Method 1.** [It basically never works because it depends on the presence of a *benign lecturer* who tells you how to factor the characteristic polynomial, while in practice it's very hard to know all the roots of which since the matrix could be huge. Of course when it works it works well.]

**Example 2.2.1.** By some divine inspiration we know that $c_A = (x-2)(x-3)^3$ where

$$A = \begin{pmatrix} 4 & 0 & -1 & -1 \\ 1 & 2 & 0 & 0 \\ 2 & -2 & 2 & -2 \\ -1 & 1 & 0 & 3 \end{pmatrix}.$$

$\mu_A$ has to have roots $2, 3$ and divide $c_A$, i.e.

$$\mu_A = \begin{cases} (x-2)(x-3) \\ (x-2)(x-3)^2 \\ (x-2)(x-3)^3 \end{cases},$$

and it turns out that it's the third one since $(A - 2I_4)(A - 3I_4), (A - 2I_4)(A - 3I_4)^2 \neq 0$.

**Method 2.** which also makes sense of $\mu_T$ of a linear self-map $T : V \to V$, $\dim V = n$ (and $\mu_T = \mu_A$ for any $A$ representing $T$ with respect to some basis) and depends on

**Lemma 2.2.2.** Given linear map $T : V \to V$ where $V$ is $n$-dimensional, assume $W_1, W_2, \ldots, W_k \subseteq V$ are finitely many $T$-invariant subspaces that span $V$ and $T = W_1 + W_2 + \cdots W_k$. (not necessarily direct sum). Then

$$\mu_T = \text{lcm } (\mu_1, \mu_2, \ldots, \mu_k)$$

where $\mu_i$ is the minimal polynomial of $T|_{W_i} : W_i \to W_i$.

*Proof.* Let $f = \text{lcm } (\mu_1, \mu_2, \ldots, \mu_k)$. We will prove $f|\mu_T$ and $\mu_T|f$. We write $f(x) = g_i(x)\mu_i(x)$. Let $v \in W_i$, then

$$f(T)v = g_i(T)\mu_i(T)v = g_i\left(T|_{W_i}\right) \underbrace{\mu_i\left(T|_{W_i}\right)}_{0 \text{ by definition of } \mu_i} v = 0,$$

but since $W_i$'s span $V$ we have $f(T)v = 0\ \forall v \in V$, i.e. $f(T) = 0 \Rightarrow \mu_T|f$.
On the other hand, $\mu_T(T) = 0$, so if $v \in W_i$ then $\mu_T(T)v = \mu_T\left(T|_{W_i}\right) v = 0 \Rightarrow \mu_T\left(T|_{W_i}\right) = 0 \Rightarrow \forall i, \mu_i|\mu_T \Rightarrow \text{lcm } (\mu_i, \ldots \mu_k) = f(x)|\mu_T$. □

The algorithm for $T : V \to V$ works like this:
Pick any $v \in V \neq 0$ and consider $W \subset V = \text{span}\left(v, Tv, T^2v, T^3v, \ldots\right)$ which is by construction $T$-invariant. Let $d \geq 1$ be the smallest positive integer such that $v, Tv, \ldots, T^dv$ are linearly dependent. This also means $v, Tv, \ldots, T^{d-1}v$ are linearly independent and $\deg \mu_{T|_W} \geq d$ since if $p$ has $\deg \leq d-1$ then $p(T)v \neq 0$. In particular this means $\exists$ a nontrivial linear dependency relation

$$T^dv + c_{d-1}T^{d-1}v + \cdots + c_1Tv + c_0v = 0,$$

the leading coefficient is nonzero by the linear independency of the $d-1$ vectors after. Then

$$x^d + c_{d-1}x^{d-1} + \cdots + c_1x + c_0,$$

is the minimal polynomial $\mu_{T|_W}(x)$. The algorithm proceeds as follows: put $W_1 = W, v_1 = v$.

- If $W_1 = V$, we are done.
- If $W_1 \subsetneqq V$, pick any $v_2 \in V \backslash W_1$ and proceed as before.

## 2.3 Jordan canonical form

Now recall Definition 3 and

**Definition 2.3.1.** A *Jordan chain* of length $k$ associated with eigenvalue $\lambda$ is an ordered $k$-tuple $(v_1, v_2, \ldots, v_k)$ where $v_i \in V$ with the following properties:

1. $0 \neq v_k \in N_k(T, \lambda) \backslash N_{k-1}(T, \lambda)$
2. $(T - \lambda\ \text{id}_V)v_k = v_{k-1}$
   $(T - \lambda\ \text{id}_V)v_{k-1} = v_{k-2} \quad \cdots$
   $(T - \lambda\ \text{id}_V)v_2 = v_1$
   $(T - \lambda\ \text{id}_V)v_1 = (T - \lambda\ \text{id}_V)^k v_k = 0$
3. $v_i$'s are nonzero

Later we'll prove that $v_i$'s are linearly independent.

Subspace $W = \text{span}\,(v_1, \ldots, v_k) \subset V$ is $T$-invariant. Question is what is the matrix of $T|_W : W \to W$ with respect to the basis $(v_1, \ldots, v_k)$? From the properties above, $v_1$ is just the usual eigenvector so $Tv_1 = \lambda v_1$, and $v_2 = v_1 + \lambda v_2$, $v_3 = v_2 + \lambda v_3$ and so on. We hence write the $k \times k$ matrix

$$\begin{pmatrix} \lambda & 1 & 0 & & & 0 \\ 0 & \lambda & 1 & & & \\ 0 & 0 & \lambda & 1 & & \\ 0 & 0 & 0 & \ddots & \ddots & \\ \vdots & \vdots & \vdots & & & 1 \\ 0 & 0 & 0 & & & \lambda \end{pmatrix},$$

with $\lambda$ on the diagonal, 1 on the superdiagonal and 0 elsewhere.

**Definition 2.3.2.** The form described above called a *Jordan block* of degree $k$, denoted $J_{\lambda,k}$.

Now this is nice, but you can't expect an entire vector space to have a basis consisting of just a single Jordan chain; you put multiple together to form what's called Jordan basis:

**Definition 2.3.3.** A *Jordan basis* for $V$ (associated with $T$) is a basis that is a union of finitely many Jordan chains.

**Theorem 2.3.4** (Jordan canonical form)**.** Let $V$ be an $n$-dimensional vector space over $\mathbb{C}$ or algebraically closed field, let $T : V \to V$ be a linear self-map. Then

1. There exists a Jordan basis for $T$. This means if we order vectors in such a basis appropriately, then the matrix of $T$ will have a block diagonal form:

$$\begin{pmatrix} J_{\lambda_1,k_1} & & & & 0 \\ & J_{\lambda_2,k_2} & & & \\ & & J_{\lambda_3,k_3} & & \\ & & & & \ddots \\ 0 & & & & \end{pmatrix}.$$

   (where the blocks can be rearranged of course).

2. The number of Jordan blocks for eigenvalue $\lambda$ and of degree at least $i$ ($i \geq 1$) is equal to

$$\dim N_i(T, \lambda) - \dim N_{i-1}(T, \lambda),$$

   which means it is uniquely determined by $T$.

3. We can read off the characteristic and minimal polynomials of the Jordan canonical form:

$$c_T(x) = (-1)^n \prod (x - \lambda_i)^{a_i}$$

   and

$$\mu_T(x) = \prod (x - \lambda_i)^{b_i}$$

   where $\lambda_i$ an eigenvalue of $T$, $a_i$ = sum of degrees of all Jordan blocks for $\lambda_i$ and $b_i$ = largest of the degrees. In particular, $T$ is diagonalisable iff $\mu_T$ does not have multiple roots, i.e. all $b_i = 1$.

*Proof.* 1. By induction on $\dim V = n$.

When $n = 1$ there's nothing to prove, $T$ is just scalar multiplication.

Now suppose statement is true for $\dim < n$. Take an eigenvalue $\lambda$ of $T$. Consider $U = \mathrm{im}(T - \lambda \mathrm{id}_V)$ where $T - \lambda \mathrm{id}_V : V \to V$. Since the map has a nontrivial kernel, $\dim U = m < \dim V$ by rank-nullity theorem. Also $U$ is $T$-invariant, i.e. $u \in U \Rightarrow T(u) \in U$: let $u = (T - \lambda \mathrm{id}_V)(u')$ where $u' \in V$, then $T(u) = T \circ (T - \lambda \mathrm{id}_V)(u') = (T - \lambda \mathrm{id}_V)(T(u')) \in U$.

Now consider $T_U = T|_U : U \to U$. By induction hypothesis, there is a Jordan basis $e_1, \ldots, e_m$ for $U$ and $T_U$.

Suppose there is $l$ Jordan chains associated with eigenvalue $\lambda$ among this basis. We want to extend these chains to chains for $T$. Suppose $w_1, \ldots, w_l$ are vectors in $V$ mapping under $T - \lambda \mathrm{id}_V$ to the last vectors in each of these $l$ Jordan chains.

Consider Ker $(T - \lambda \mathrm{id}_V)$ whose dimension is $n - m$ again by R-N theorem, i.e. eigenspace for $\lambda$ is $(n - m)$-dimensional and each of the first vectors in the $l$ Jordan chains above is in this eigenspace, since they are eigenvectors instead of generalised eigenvectors. So they span an $l$-dimensional subspace of the eigenspace associated with $\lambda$. Pick

$$\underbrace{w_{l+1}, \ldots, w_{n-m}}_{n-m-l \text{ eigenvectors that are not in } U, \text{ hence are Jordan chains of length 1}} \in V$$

completing the $l$ eigenvectors for $\lambda$ to be a basis for the eigenspace for $\lambda$. We claim

$$\underbrace{w_1, \ldots, w_l, w_{l+1}, \ldots, w_{n-m}, e_1, \ldots, e_m}_{n \text{ vectors which is a disjoint union of Jordan chains}}$$

are linearly independent, i.e.

$$\alpha_1 w_1 + \cdots + \alpha_l w_l + \alpha_{l+1} w_{l+1} + \cdots \alpha_{n-m} w_{n-m} + x = 0 \qquad (*)$$

iff $\forall i, \alpha_i = 0$ where $x \in U$. Apply $T - \lambda \mathrm{id}_V$ to $(*)$. The $n - m - l$ vectors in the middle are annihilated since they are eigenvectors. We then have

$$\alpha_1 \underbrace{(T - \lambda \mathrm{id}_V)(w_1)}_{\text{last vector of the first Jordan chain } e_{i_1}} + \cdots + \alpha_l \underbrace{(T - \lambda \mathrm{id}_V)(w_l)}_{\text{of the } l\text{th chain } e_{i_l}} + \underbrace{(T - \lambda \mathrm{id}_V)(x)}_{\text{linear combination of } e_1, \ldots, e_m \setminus e_{i_1}, \ldots, e_{i_l}} = 0.$$

It follows that $\alpha_1 = \cdots = \alpha_l = 0$ and $(T - \lambda \mathrm{id}_V)(x) = 0$, i.e. $x$ is an eigenvector for $\lambda$. We therefore have a Jordan basis for $T$.

2. If a matrix has block diagonal form, then dimension of kernel of matrix is sum of dimensions of kernels of blocks. So it suffices to prove for a single Jordan block $A = J_{\lambda,k}$. Consider $\dim \mathrm{Ker}\ (A - \lambda I_k)^i$.

| $i$ | 0 | 1 | 2 | $\cdots$ | $k-2$ | $k-1$ | $k$ | $k+1$ | $k+2$ | $\cdots$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\dim \mathrm{Ker}\ (A - \lambda I_k)^i$ | 0 | 1 | 2 | $\cdots$ | $k-2$ | $k-1$ | $k$ | $k$ | $k$ | $\cdots$ |

Hence the number of Jordan blocks of precise degree $i$/size $i$ occurring in the Jordan canonical form for $T$ is

$$\underbrace{(\dim N_i(T, \lambda) - \dim N_{i-1}(T, \lambda))}_{\geq i} - \underbrace{(\dim N_{i+1}(T, \lambda) - \dim N_i(T, \lambda))}_{\geq i+1}$$
$$= 2 \dim N_i(\lambda, T) - \dim N_{i-1}(T, \lambda) - \dim N_{i+1}(T, \lambda).$$

3. Observe that if $M$ has block diagonal form

$$\begin{pmatrix} M_1 & & & 0 \\ & M_2 & & \\ & & \ddots & \\ 0 & & & M_r \end{pmatrix}$$

where $M_i$ are some square matrices, then

$$c_M = \prod_i c_{M_i}$$

and

$$\mu_M = \text{lcm } (\mu_{M_1}, \ldots, \mu_{M_i}).$$

The desired result follows immediately follows since $\mu_{J_{\lambda,k}}(x) = (x-\lambda)^k$ and $c_{J_{\lambda,k}}(x) = (-1)^k(x-\lambda)^k$.

□

**Remark.** If $\lambda_1, \ldots, \lambda_r$ are eigenvalues of $T : V \to V$, then $\forall i$,

$$N_{i_1}(T, \lambda_j) \cap N_{i_2}(T, \lambda_k) = \{0\}$$

where $\lambda_i \neq \lambda_j$. More generally, if $\lambda_1, \ldots, \lambda_r$ are distinct, then the sum of the full generalised eigenspace associated with $\lambda_1, \ldots, \lambda_r$ is direct.

$$\begin{pmatrix} J_{\lambda,1} & \\ & J_{\lambda,2} \end{pmatrix}$$

Suppose $A$ is a matrix in Jordan canonical form representing $T$ with respect to some Jordan basis.

$$N_i(T, \lambda) = \text{Ker } (T - \lambda \text{id}_V)^i$$
$$N_1(T, \lambda) \subseteq N_2(T, \lambda) \subseteq \cdots$$
$$N_\infty(T, \lambda) = \bigcup_i N_i(T, \lambda)$$

whose elements are generalised eigenvectors and spanned by the first $n_1$ vectors in the Jordan basis and $N_\infty(T, \mu)$ is spanned by he first $n_2$ vectors in the Jordan basis, etc.
$p \in [x]$, $T : V \to V$,

$$N_p := \text{Ker } (p(T)).$$

Then if $p, q$ relatively prime,

$$N_p \cap N_q = \{0\},$$

i.e. $\exists r, s \in [x] : rp + sq = 1$. If $v \in N_p \cap N_q$ then

$$(r(T)p(T) + s(T)q(T))v = v,$$

and just apply this to $(x-\lambda)^i$ and $(x-\mu)^j$ for $\lambda, \mu$ eigenvalues of $T$.

If $A$ is a $2 \times 2$ matrix with $c_A(x) = \mu_A(x) = (x-\lambda)^2$, then the Jordan canonical form of $A$ is $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ since the maximum degree of the block is 2, the same size with the whole matrix, which is therefore a Jordan block itself. But if $\mu_A(x) = x - \lambda$ then the Jordan canonical form is $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$, i.e. diagonalisable since $\mu_A$ has no multiple roots.

## 2.4 General brute-force algorithm for computing the Jordan canonical form $J$ for a matrix $A \in \mathbb{C}^{n\times n}$ and an invertible matrix $P$ such that $P^{-1}AP = J$ if you know all the eigenvalues.

1. Computing $J$ is straightforward. For an eigenvalue $\lambda$ of $A$, the number of Jordan blocks of size/degree $i$ in $J$ is $2\dim N_i(\lambda, T) - \dim N_{i-1}(T,\lambda) - \dim N_{i+1}(T,\lambda)$.

   Sometimes there are shortcuts for "small" $n$ using $c_A$ and $\mu_A$, but note that $\begin{pmatrix} J_{\lambda,3} & & \\ & J_{\lambda,3} & \\ & & J_{\lambda,1} \end{pmatrix}$ and $\begin{pmatrix} J_{\lambda,3} & & \\ & J_{\lambda,2} & \\ & & J_{\lambda,2} \end{pmatrix}$ share the same polynomials $c(x) = -(x-\lambda)^7$ and $\mu(x) = (x-\lambda)^3$ but their form are different.

2. Computing $P$ is harder in general. Pick an eigenvalue $\lambda$, there are a number of Jordan blocks for $\lambda$ of degrees, say,
$$N_1 \geq N_2 \geq \cdots \geq N_r \in \mathbb{N}.$$
   Pick $v \in V : (A-\lambda I)^{N_1} v = 0$ but $(A-\lambda I_n)^{N_1-1} v \neq 0$. Put
$$v_{1,1} = v, v_{1,2} = (A-\lambda I_n)v, v_{1,3} = (A-\lambda I_n)^2 v, \ldots, v_{1,N_1} = (A-\lambda I_n)^{N_1-1} v,$$
   a Jordan chain of length $N_1 - 1$. If there's no $N_2$ we're done. If not, pick $v_{2,1} \in V$ : $(A-\lambda I_n)^{N_2} v_{2,1} = 0$, but
$$(A-\lambda I_n)^{N_2-1} v_{2,1} \notin \operatorname{span}\{v_{1,1}, v_{1,2}, \ldots, v_{1,N_1}\}.$$
   Then put $v_{2,2} = (A-\lambda I_n)v_{2,1}, \ldots, v_{2,N_2} = (A-\lambda I_n)^{N_2-1} v_{2,1}$. If there's no $N_3$ we're done. If not, pick $v_{3,1} \in V : (A-\lambda I_n)^{N_3} v_{3,1} = 0$ but
$$(A-\lambda I_n)^{N_3-1} v_{3,1} \notin \operatorname{span}\{v_{1,1}, \ldots, v_{1,N_1}, v_{2,1}, \ldots, v_{2,N_2}\},$$
   and proceed similarly.

   Once we're done with $\lambda$, write the list of vectors produced in reverse order and continue with the next eigenvalue unless there are none left. Write these vectors into the columns of a matrix $P$. Then $P^{-1}AP = J$.

The point of all this is: it can be done!

**Example 2.4.1.**
$$A = \begin{pmatrix} -1 & -3 & -1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 1 & -1 \end{pmatrix}.$$

Then $c_A = (-1-x)^2(2-x)^2$, i.e. eigenvalues $-1$ and 2. Compute the dimensions of eigenspaces. For $\lambda = -1$:
$$A + I_4 = \begin{pmatrix} 0 & -3 & -1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 3 & 1 & 0 \end{pmatrix}$$

which has a 2-dim column space and therefore 2-dim null space by rank-nullity theorem. We further have the two eigenvectors $v_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, v_2 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$ which is in Ker $(A - 2I_4)$ and nonzero; and for $\lambda = 2$:

$$A - 2I_4 = \begin{pmatrix}-3 & -3 & -1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 3 & 1 & 3\end{pmatrix}$$

which has a 3-dim column space and therefore 1-dim null space. $\Rightarrow$

$$J = \begin{pmatrix}-1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 2\end{pmatrix}.$$

To find $P$, find a vector $v_4 \in$ Ker $(A-2I_4)^2$ but not in Ker $(A-2I_4)$, and then $v_3 = (A-2I_4)v_4$, then put $v_1, v_2, v_3, v_4$ in columns:

$$P = \begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 1 & \underbrace{1}_{v_3} & \underbrace{0}_{v_4}\end{pmatrix}.$$

**Example 2.4.2.**

$$A = \begin{pmatrix}3 & 2 & 1\\ 0 & 3 & 1\\ -1 & -4 & -1\end{pmatrix},$$

then $c_A = (2-x)^2(1-x)$ and $\mu_A = (x-2)^2(x-1)$ (can be acquired by Cayley–Hamilton theorem and check the two possibilities). Hence

$$J = \begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{pmatrix}$$

Then we compute eigenvectors for $\lambda = 1$, which gives $v_3 = \begin{pmatrix}0\\1\\-2\end{pmatrix}$. Then find $v_2 \in$ Ker $(A-2I_3)^2$ but not in Ker $(A - 2I_3)$, which gives, e.g. $v_2 = \begin{pmatrix}2\\0\\-1\end{pmatrix}$, and $v_1 = (A - 2I_3)v_2 = \begin{pmatrix}1\\-1\\-1\end{pmatrix}$. Then

$$P = \begin{pmatrix}1 & 2 & 0\\ -1 & 0 & 1\\ 1 & 1 & -2\end{pmatrix}.$$

# 3 Functions of matrices

Given a matrix $A \in \mathbb{C}^{n\times n}$, how do you compute $A^n$ for $n$ very large? $n = 2022$ maybe?

Where does that question occur naturally? It comes from theory of dynamical system where you may be interested in the behaviour of solutions to an ODE of the following form

$$\dot{x} = f(x), \quad x(t) \in \mathbb{R}^n \quad f \in C(\mathbb{R}^n, \mathbb{R}^n)$$

locally near an equilibrium (i.e. constant solution). $f$ may have a zero at $x \in \mathbb{R}^n$, $x(t) = x_0\ \forall t$. For this one linearises the equation near $x_0$ and choose coordinates such that $x_0 = 0$, giving $\dot{x} = Ax$ where $A \in \mathbb{R}^{n\times n}$, actually called the Jacobian. Then you discretise time (e.g. certain clicks of clock) which give us the problem $x(k+1) = Ax(k)$ which recursively defines vectors $x(0), x(1), \ldots \in \mathbb{R}^n$. For instance we can look at Fibonacci sequence, which is recursively defined, in this manner.

This is a different part of mathematics which we'll not get into.

So back to the algebraic question. If we know the Jordan canonical form $J$ for $A$ the question is easier. Then $A = PJP^{-1}$, so

$$A^n = \underbrace{(PJP^{-1})(PJP^{-1})\cdots(PJP^{-1})}_{n \text{ times}} = PJ^nP^{-1}.$$

So if $J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix}$ then $J^n = \begin{pmatrix} J_1^n & & & \\ & J_2^n & & \\ & & \ddots & \\ & & & J_r^n \end{pmatrix}$. Easy, we just need to compute $J_{\lambda,k}^n$ for single Jordan block $J_{\lambda,k}$.

**Example 3.0.1.** We have $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}$ and $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^3 = \begin{pmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{pmatrix}$ and hence by induction we can prove that $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}$.

Also $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & \lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix}$.

So the general answer for $k \times k$ matrices is:

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ & \lambda^n & n\lambda^{n-1} & \binom{n}{1}\lambda^{n-1} & \ddots & \\ & & \lambda^n & n\lambda^{n-1} & \binom{n}{1}\lambda^{n-1} & \ddots \\ & & & \ddots & n\lambda^{n-1} & \binom{n}{1}\lambda^{n-1} \\ & & & & \lambda^n & n\lambda^{n-1} \\ 0 & & & & & \lambda^n \end{pmatrix}$$

i.e. you get $\lambda^n$ on the diagonal and $\binom{n}{i}\lambda^{n-i}$ on the $i$th diagonal above it.

*Handwavy proof.* You could use induction, or consider

$$v_{i+1} = \frac{x^i}{i!}e^{\lambda x}, \quad i = 0, \ldots, k-1$$

which are $k$ smooth functions on $\mathbb{R}$. Now consider the vector subspace of the space of smooth functions $V = \operatorname{span}\{v_1, \dots, v_k\}$ and look at the derivative map $\frac{\mathrm{d}}{\mathrm{d}x} : V \to V$ and the matrix of this with respect to $(v_1, \dots, v_k)$ is $J_{\lambda,k}$. So to compute $J^n_{\lambda,k}$, we have to figure out what $\dfrac{\mathrm{d}^n}{\mathrm{d}x^n}$ does to $\dfrac{x^i}{i!}e^{\lambda x}$. We can use the generalised (binomial) product rule:

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(fg) = \sum_{j=0}^{n} \binom{n}{j} \frac{\mathrm{d}^j}{\mathrm{d}x^j}(f) \frac{\mathrm{d}^{n-j}}{\mathrm{d}x^{n-j}}(g),$$

which gives us entries of the matrix. □

There's a second method of computing high powers of matrices. Choose a polynomial $\Psi(x) \in \mathbb{C}[x]$ that satisfies $\Psi(A) = 0$, i.e. a multiple of minimal polynomial $\mu_A(x)$ (e.g. $\mu_A(x)$ itself or $c_A(x)$. Then divide $f(x) = x^n$ by $\Psi(x)$ with remainder $h(x)$, i.e.

$$f(x) = q(x)\Psi(x) + h(x)$$

where $\deg h < \deg \Psi$ if $h \neq 0$. Then $f(A) = \underbrace{q(A)\Psi(A)}_{0} + h(A) = h(A) = A^n$. But the division is generally hard. An easy method to find $f(x)$ is Lagrange interpolation. Suppose $\Psi(x) = \prod (x - \alpha_i)^{m_i}$, then $f$ have the same value and $(m_j - 1)$th derivatives at $\alpha_j$ as $h$:

$$f^{(t)}(\alpha_j) = h^{(t)}(\alpha_j), \quad j = 1, \dots, k, \quad t = 1, \dots, m_j - 1.$$

These are $m_1 + \cdots m_k$ equations for the $m_1 + \cdots + m_k$ unknowns coefficients of $h$: solvable!

**Example 3.0.2.** We know $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & -2 & -2 \end{pmatrix}$, $J = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$ and $J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -2 & 0 \end{pmatrix}$. So $J = P^{-1}AP$ and $\mu_A(x) = (x+2)^2$. We have $A^n = PJ^nP^{-1}$ and $J^n = \begin{pmatrix} (-2)^n & n(-2)^{n-1} & 0 & 0 \\ 0 & (-2)^n & 0 & 0 \\ 0 & 0 & (-2)^n & n(-2)^{n-1} \\ 0 & 0 & 0 & (-2)^n \end{pmatrix}$ and the rest is just computation.

Lagrange interpolation is easier here. Let $\Psi(x) = \mu_A(x) = (x+2)^2$. So choose a lower degree remainder $h(x) = \alpha x + \beta$ which is linear, such that

$$h(-2) = -2\alpha + \beta = (-2)^n$$

and

$$h'(-2) = \alpha = n(-2)^{n-1}.$$

So

$$h = n(-2)^{n-1}x + (1-n)(-2)^n,$$

and therefore

$$A^n = h(A) = n(-2)^{n-1}A + (1-n)2^n \cdot I_4$$

which gives same answer almost immediately, without need to compute Jordan canonical form. [JCF is still significant theoretically!]

**Example 3.0.3** (Fibonacci numbers)**.** $F_0 = 0,\ F_1 = 1,\ F_n = F_{n-1} + F_{n-2},\ n \geq 2.$
We then have

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix}.$$

So

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n}_{A^n} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The characteristic polynomial

$$c_A(x) = (-x)(1-x) - 1 = x^2 - x - 1$$

gives roots

$$\lambda = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2} = 1 - \lambda$$

which are eigenvalues. (Golden ratio)
Use interpolation to compute $A^n$. Choose $\Psi(x) = c_A(x) = \mu_A(x) = x^2 - x - 1$ (diagonalisable so equal) such that $\Psi(A) = 0$. We want to write

$$x^n = c_A(x) \cdot q(x) + \underbrace{h(x)}_{\text{remainder of degree } \leq 1}.$$

So $h(x) = \alpha x + \beta$ such that

$$h(\lambda) = \alpha\lambda + \beta = \lambda^n$$

and

$$h(1-\lambda) = \alpha(1-\lambda) + \beta = (1-\lambda)^n$$

which gives

$$\begin{cases} \alpha = \dfrac{\overbrace{\lambda^n - (1-\lambda)^n}^{\mu_n}}{\sqrt{5}} \\ \beta = \dfrac{\overbrace{\lambda^{n-1} - (1-\lambda)^{n-1}}^{\mu_{n-1}}}{\sqrt{5}} \end{cases}.$$

So

$$A^n = h(A) = \frac{\mu_n}{\sqrt{5}} A + \frac{\mu_{n-1}}{\sqrt{5}} I_2 = \begin{pmatrix} 0 & \frac{\mu_n}{\sqrt{5}} \\ \frac{\mu_n}{\sqrt{5}} & \frac{\mu_n}{\sqrt{5}} \end{pmatrix} + \begin{pmatrix} \frac{\mu_{n-1}}{\sqrt{5}} & 0 \\ 0 & \frac{\mu_{n-1}}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{\mu_{n-1}}{\sqrt{5}} & \frac{\mu_n}{\sqrt{5}} \\ \frac{\mu_n}{\sqrt{5}} & \frac{\mu_n + \mu_{n-1}}{\sqrt{5}} \end{pmatrix}$$

and we conclude that $F_n = \dfrac{\mu_n}{\sqrt{5}}$.

## 3.1 Solving systems of linear differential equations with constant coefficients using exponential function of a matrix

$$v : \mathbb{R} \to \mathbb{R}^n,\ v(t) = \begin{pmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{pmatrix}$$

which is smooth satisfies

$$\frac{d}{dt}v = \dot{v} = Av, \quad A \in \mathbb{R}^{n\times n}, \quad v(0) = v_0 \in \mathbb{R}^n,$$

i.e.

$$\begin{pmatrix} \dot{v}_1(t) \\ \vdots \\ \dot{v}_n(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{pmatrix}.$$

e.g.

$$\dot{v}_1 = -2v_1 + 3v_2$$
$$\dot{v}_2 = 4v_1 + e^{2\pi+e+\sqrt{2}\cdot}$$

If $n = 1$ solution is simply $v(t) = e^{tA} \cdot v_0$ where $A$ is just a real number. The hope is this is also the general solution once we make sense of powers of matrices $e^{tA}$, which is a $n \times n$ matrix whose entries smoothly depend on $t$.

$t$ is not much of a problem so we need to define $e^A$ for $A \in \mathbb{C}^{n\times n}$ in general. There are at least 2 approaches, which turned out to be equivalent.

1. We know $e^x$ is defined by power series $\sum \frac{x^n}{n!}$ and we know how to do polynomials of matrices, so we already know how to make sense of partial sums (which are $n\times n$ matrices). Standard guess is then
$$e^A = I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots \qquad (*)$$
and hoping the limit converges. We can use the norm
$$\|M\| = \sup \|Mv\|_2$$
to know the answer of that (which is yes), but it's not of the interest of this module. Suppose $J = P^{-1}AP$ is a Jordan canonical form of $A$. Then $A = PJP^{-1}$, so $Pe^JP^{-1}$ by substituting into $(*)$, which is easier since powers of Jordan canonical form are easier. Suppose
$$J = J_{\lambda_1,k_1} \oplus J_{\lambda_2,k_2} \oplus \cdots \oplus J_{\lambda_t,k_t},$$
then
$$e^J = e^{J_{\lambda_1,k_1}} \oplus e^{J_{\lambda_2,k_2}} \oplus \cdots \oplus e^{J_{\lambda_t,k_t}}.$$
So we focus on single Jordan block
$$A = J_{\lambda,k} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda \end{pmatrix}.$$
We already know $f(A)$ where $f$ is a power of $x$, $f(x) = x^N$:
$$f(A) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \cdots & \cdots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ & f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \ddots & \\ & & f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \ddots \\ & & & \ddots & f'(\lambda) & \frac{f''(\lambda)}{2} \\ & & & & f(\lambda) & f'(\lambda) \\ 0 & & & & & f(\lambda) \end{pmatrix} \qquad (**)$$

and

$$f(J) = f(J_{\lambda_1,k_1}) \oplus \cdots \oplus f(J_{\lambda_t,k_t}).$$

We could use $(**)$ to define $f(A)$ for any reasonably nice function $f$ (smooth at least).

2. Write $A = PJP^{-1}$ where $J$ is a Jordan canonical form. Define

$$f(A) = Pf(J)P^{-1}$$

where

$$f(J) := f(J_{\lambda_1,k_1}) \oplus \cdots \oplus f(J_{\lambda_t,k_t})$$

and finally define $f(J_{\lambda,k})$ by $(**)$.

This agrees with $e^A$ for $f(x) = e^x$ as defined previously using limit of partial sums.

Caveat: some rules such as $e^{B+C} = e^B e^C$ are not valid in general where $B, C$ are matrices. If they commute then valid.

**Example 3.1.1.**

$$A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & -2 & -2 \end{pmatrix},$$

$$c_A(x) = (-2-x)^4, \quad \mu_A(x) = (x+2)^2.$$

Let's use interpolation to compute $e^A$. We have

$$f(x) = q(x)\Psi(x) + h(x)$$

where we choose $\Psi(x) = \mu_A(x)$ let $f(x) = e^x$, the function we are interested in.

(We can actually divide power series by polynomial, which still gives a polynomial remainder of degree less than polynomial dividing. $q(x)$ would be a power series.)

$h(x)$ then can be chosen to be linear: $h(x) = \alpha x + \beta$ such as

$$h(-2) = -2\alpha + \beta = e^{-2}$$

and

$$h'(-2) = \alpha = e^{-2},$$

which give us

$$\begin{cases} \alpha = e^{-2} \\ \beta = 3e^{-2} \end{cases}$$

and therefore

$$e^A = f(A) = h(A) = e^{-2}A + 3e^{-2}I_4 = \begin{pmatrix} e^{-2} & 0 & 0 & 0 \\ 0 & e^{-2} & e^{-2} & 0 \\ 0 & 0 & e^{-2} & 0 \\ e^{-2} & 0 & -2e^{-2} & e^{-2} \end{pmatrix}$$

which is much simpler computing involving change of basis matrix.

**Example 3.1.2** (Harmonic oscillator)**.**

$$y''(t) = -y(t)$$

We look at the vector

$$x(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}$$

to rewrite the equation to system of linear equations

$$x'(t) = \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{A} x(t).$$

We need to specify both $y(0), y'(0)$ to get a well-defined initial value problem. Now compute $e^{tA}$. Since $c_A(x) = \mu_A(x) = x^2 + 1$, eigenvalues of $A$ (and $tA$ for $t$ fixed) are $\pm i \in \mathbb{C}$. Let $f(x) = e^{tx}$. Again $h(x) = \alpha x + \beta$ such as

$$h(i) = \alpha i + \beta = e^{ti}$$

and

$$h(-i) = -\alpha i + \beta = e^{-ti}$$

which give us

$$\begin{cases} \alpha = \dfrac{e^{ti} - e^{-ti}}{2i} = \sin t \\ \beta = \dfrac{e^{ti} + e^{-ti}}{2} = \cos t \end{cases}$$

So $e^{tA} = \sin t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \cos t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$. Therefore

$$\begin{pmatrix} y(t) \\ y'(t) \end{pmatrix} = x(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix}$$

$\Rightarrow y(t) = \cos t \cdot y(0) + \sin t \cdot y'(0)$.

# 4 Bilinear maps and quadratic forms

**Definition 4.0.1.** $V, W$ vector spaces over $K$. A *bilinear map* on $(V, W)$ is a map

$$\tau : V \times W \to K$$

such that for arbitrary (fixed) $v \in V$ the map $\tau(v, -) : W \to K$ is linear, and for arbitrary (fixed) $w \in W$, $\tau(-, w) : V \to K$ is linear.

e.g. scalar product on $\mathbb{R}^n$

Suppose $E = (e_1, \ldots, e_n)$ and $F = (f_1, \ldots, f_m)$ are ordered bases for $V, W$. Then we write

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in V, \quad w = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in W$$

where $v = \sum x_i e_i,\ w = \sum y_i f_i$. So

$$\tau(v,w) = \tau\left(\sum x_i e_i, \sum y_i f_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i \tau(e_i, f_j) y_j.$$

And if we define $\alpha_{ij} := \tau(e_i, f_j)$ and $A := (\alpha_{ij})$ which is a $m \times n$ matrix, then

$$\tau(v,w) = v^T A w.$$

We conclude that bilinear maps correspond to $n \times m$ matrices $A$ after choice of ordered bases.
What happens to $A$ if we change bases in $V, W$?
Choose $E' = (e'_1, \ldots, e'_n),\ F' = (f_1, \ldots, f'_m)$. We have the change of basis matrix

$$P = \mathcal{M}(\mathrm{id}_V)^E_{E'} \qquad \text{write } E' \text{ in columns in terms of } E$$

and

$$Q = \mathcal{M}(\mathrm{id}_V)^F_{F'},$$

then $Pv_{E'} = v_E$ and $Qw_{F'} = w_F$. So

$$\tau(v,w) = v_E^T A w_F = \left(Pv_{E'}\right)^T A \left(Qw_{F'}\right) = v_{E'}^T \left(P^T A Q\right) w_{F'},$$

where $P^T A Q$ is the matrix of $\tau$ with respect to the new bases.
Observe that $P, Q$ are invertible, so rank of matrix representing $\tau$ is independent of choice of bases. (Justifying the notion of rank of bilinear map $\tau$.)

**Definition 4.0.2.** If $V = W$ (which happens frequently), we call such bilinear map $\tau$ a *bilinear form* on $V$.

Compare with linear map − operator.
If we change basis in $V$ using $P$, the matrix $A$ representing the bilinear form changes to

$$P^T A P.$$

Such $A$ are called congruent.

**Definition 4.0.3.** Let $\tau : V \times W \to K$ be a bilinear map. Consider

$$\{w \in W : \tau(v,w) = 0\ \forall v \in V\}.$$

This is called the right *radical* of $\tau$. Left radical is defined similarly:

$$\{v \in V : \tau(v,w) = 0\ \forall w \in W\}.$$

In terms of the matrix, since $\tau(v,w) = v_E\left(Aw_F\right) = \left(A^T v_E\right)^T w_F$, right radical is simply kernel of $A$, and left radical is kernel of $A^T$. They have the same dimension by rank-nullity theorem. (row rank = column rank) In particular if $V = W$ and $\dim V = n$ and rank $\tau$ = rank $A = r$, then $\dim \mathrm{Ker}\ A = \dim \mathrm{Ker}\ A^T = n - r$.

We now want to classify bilinear forms since they are too general to study.

**Definition 4.0.4.** Let $\tau : V \times V \to K$ be a bilinear form. We call $\tau$

- *symmetric* if $\tau(v,w) = \tau(w,v)\ \forall v, w \in V$.
- *anti-symmetric* (or *alternating/skew-symmetric*) if $\tau(v,v) = 0\ \forall v \in V$. This implies

  $\tau(v_1+v_2, v_1+v_2) = 0 = \tau(v_1,v_1) + \tau(v_2,v_2) - \tau(v_1,v_2) + \tau(v_2,v_1) = -\tau(v_1,v_2) + \tau(v_2,v_1)$

  i.e. $\tau(v_1,v_2) = -\tau(v_2,v_1)$ which does not imply back unless $1+1=0$.

  - Observation: $\tau$ is symmetric/anti-symmetric if and only if representing matrix $A$ is symmetric/anti-symmetric (latter means zeros along diagonal)

**Proposition 4.0.5.** Suppose $2 \neq 0$ in $K$. Then any bilinear form $\tau : V \times V \to K$ can be written uniquely as sum

$$\tau = \tau^s + \tau^a$$

where $\tau^s$ is symmetric and $\tau^a$ is anti-symmetric.

*Proof.* Define symmetrisation and anti-symmetrisation operators

$$\mathcal{S}(\tau)(v,w) := \frac{\tau(v,w) + \tau(w,v)}{2},$$

$$\mathcal{A}(\tau)(v,w) := \frac{\tau(v,w) - \tau(w,v)}{2},$$

then $\mathcal{S}(\tau)$ itself is symmetric and $\mathcal{A}(t)$ is anti-symmetric, and

$$\tau = \mathcal{S}(\tau) + \mathcal{A}(\tau).$$

If $\tau = \tau^s + \tau^a$ is such a decomposition, then applying $\mathcal{S}$ gives

$$\mathcal{S}(\tau) = \mathcal{S}(\tau^s) + 0 = \tau^s$$

and similarly $\mathcal{A}(\tau) = \tau^a$ which proves uniqueness. □

## 4.1 Quadratic forms

**Definition 4.1.1.** $V$ is a vector space over $K$. A *quadratic form* $q$ on $V$ is a map

$$q : V \to K$$

such that

1. $q(\lambda v) = \lambda^2 q(v) \quad \forall v \in V,\ \forall \lambda \in K$
2. $\tau_q(v,w) := q(v+w) - q(v) - q(w)$ is a symmetric bilinear form
   - This means $q(x_1b_1 + x_2b_2) = q(x_1b_1) + q(x_2b_2) + \tau_q(x_1b_1, x_2b_2) = x_1^2 q(b_1) + x_2^2 q(b_2) + x_1x_2\tau_q(b_1,b_2)$

If $B = (b_1, \ldots, b_n)$ is a basis for $V$ and $v \sum_{i=1}^n x_i b_i$, then a quadratic form is just a general quadratic polynomial in the variables $x_i$:

$$q(v) = \sum_{i,j=1}^{n} c_{ij} x_i x_j \quad c_{ij} \in K.$$

e.g. for two variables

$$3x_1^2 + 5x_1x_2 + 7x_2^2$$

Given a symmetric bilinear form $\tau$, we can make a quadratic form

$$q_\tau(v) := \tau(v, v), \quad v \in V.$$

But does $q_{\tau_q}$ where $\tau_q$ is defined above equal to $q$? Unfortunately not, it's $q(2v) - q(v) - q(v) = 4q - 2q = 2q$. So similarly $\tau_{q_t} = 2\tau$, i.e. symmetric bilinear forms are not equivalent to quadratic forms if $2 = 0$ in $K$. But if $2 \neq 0$ then $\exists$ a 1-1 correspondence between the two:

$$q \rightsquigarrow \frac{1}{2}\tau_q$$

and

$$\tau \rightsquigarrow q_\tau.$$

From now on we assume $2 \neq 0$. (e.g. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_3$).

## 4.2 Nice bases

**Theorem 4.2.1.** Suppose $\tau$ is a symmetric bilinear form on $V$ with associated quadratic form $q(v) = q_\tau(v) = \tau(v, v)$. Then $\exists$ a basis $b_1, \ldots, b_n$ for $V$ such that $\tau(b_i, b_j) = \begin{cases} 0 & i \neq j \\ \beta_i & i = j \end{cases}$ i.e. the matrix of $\tau$ in this basis is diagonal $\begin{pmatrix} \beta_1 & & & 0 \\ & \beta_2 & & \\ & & \ddots & \\ 0 & & & \beta_n \end{pmatrix}$. Equivalently, every symmetric matrix $A$ is congruent to a diagonal matrix ($\exists P$ invertible such that $P^TAP$ is diagonal). And equivalently, for any $q$, $\exists$ a basis $b_1, \ldots, b_n$ such that

$$q\left(\sum x_i b_i\right) = \sum \beta_i x_i^2,$$

a simple sum of squares.

*Proof.* (by induction on $n = \dim V$)

If $\tau(v, w) \equiv 0$, $\forall v, w \in V$ there's nothing to do.

If $\tau \not\equiv 0$ then we can find a vector $b_1 : \tau(b_1, b_1) = q(b_1) \neq 0$. Define $\beta_i := \tau(b_i, b_i)$. Consider a subset

$$W = \{v \in V : \tau(v, b_1) = 0\} \quad \text{“orthogonal complement of } b \in V \text{ with respect to } \tau\text{”}$$

If we have the linear map $\tau(-, b_1) : V \to K$, $v \mapsto \tau(v, b_1)$ then $W$ is kernel of this map. We know $\tau(b_1, b_1) \neq 0$ so it's surjective. Hence $\dim W = \dim V - 1$ by rank-nullity theorem. Apply induction hypothesis to $W$ and

$$\tau|_W : W \times W \to K$$

then we're done: $\exists b_2, \ldots, b_n$ such that $\tau(b_i, b_j) = \begin{cases} 0 & i \neq j \\ \beta_i & i = j \end{cases}$, $2 \leq i, j \leq n$, so $\tau(b_i, b_j) = 0$ for $i \geq 2$. □

Basically what we've done is constructing

$$\begin{pmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}$$

and use our hypothesis to prove that $A$ is diagonal so that the whole matrix is too.

### 4.2.1 How to find such a basis for $\tau/q$? (which is not identically zero)

Suppose

$$B = (\beta_i)_{1 \le i,j \le n}$$

is the matrix for $\tau$ with respect to some initially given basis $b_1, \ldots, b_n$ which is not good enough. We want to change basis $(b_1, \ldots, b_n)$ successively until we arrive at the basis as in Theorem 4.2.1.

1. How do we find $b_1' : \tau(b_1', b_1') = q(b_1') \neq 0$? There are 2 possibilities:

   - $\exists b_i : \tau(b_i, b_i) \neq 0$ then we can just set $b_1' = b_i$.
   - The other possibility is of course $\tau(b_i, b_i) = 0 \ \forall i = 1, \ldots, n$. But there exists $i, j$ : $\tau(b_i, b_j) \neq 0$ since it's not identically zero. So put $b_1' = b_i + b_j$. Then $q(b_1') = \tau(b_1', b_1') = \tau(b_i, b_i) + \tau(b_j, b_j) + 2\tau(b_i, b_j) = 2\tau(b_i, b_j)$ which is nonzero.

   What does this mean in terms of quadratic forms? Suppose

   $$q(x_1, \ldots, x_n) = \beta_{11}x_1^2 + \cdots \beta_{nn}x_n^2 + 2\beta_{12}x_1x_2 + \cdots$$

   then we can have some $\beta_i \neq 0$, in that case we just declare that nonzero coefficient to be the new coordinate, but we can also have no squares and only mixed terms, then $\exists$ a mixed term $2\beta_{ij}x_ix_j$ such that $\beta_{ij} \neq 0$, so if we have

   $$x_ib_i + x_jb_j$$

   we can put $b_i' = b_i + b_j,\ b_j' = b_j$ then we can write it in new basis:

   $$x_i'b_i' + x_j'b_j' = (x_i)b_i' + (x_j - x_i)b_j',$$

   so $x_i' = x_i$ and $x_j' = x_j - x_i$, i.e. $x_i = x_i'$ and $x_j = x_j' + x_i'$, therefore $2\beta_{ij}x_ix_j$ is

   $$2\beta_{ij}(x_i')(x_j' + x_i')$$

   which gives a square!

2. We can assume we have a basis $b_1, \ldots, b_n$ (different from the original) such that $\tau(b_i, b_i) \neq 0$. How do we get a basis for $W$?

   We want the basis satisfy $\tau(b_1, b_i) = 0$ for $i = 2, \ldots, n$.

   We write

   $$\tau\left(b_1, b_i - \frac{\beta_{1i}}{\beta_{11}}b_1\right) = \tau(b_1, b_i) - \frac{\beta_{1i}}{\beta_{11}}\tau(b_1, b_1) = \beta_{1i} - \beta_{1i} = 0.$$

Hence we have the basis

$$b_1, b_2 - \frac{\beta_{12}}{\beta_{11}} b_1, b_3 - \frac{\beta_{13}}{\beta_{11}} b_1, \ldots, b_n - \frac{\beta_{1n}}{\beta_{11}} b_1$$

for $W$.

What does this mean in terms of quadratic forms? We have

$$q(x_1, \ldots, x_n) = \underbrace{\beta_{11}x_1^2 + 2\beta_{12}x_1x_2 + 2\beta_{13}x_1x_3 + \cdots + 2\beta_{1n}x_1x_n}_{\text{all terms involving } x_1} + C$$

where $x_i$ are new coordinates $C$ does not involve $x_1$ any more. We can now do a Babylonian trick, completing the square:

$$q(x_1, \ldots, x_n) = \beta_{11} \left( \underbrace{x_1 + \frac{\beta_{12}}{\beta_{11}} x_2 + \cdots + \frac{\beta_{1n}}{\beta_{11}} x_n}_{\text{we declare this as } x_1'} \right)^2 + C'$$

where $C'$ is some mess but also only involves $x_2, \ldots, x_n$, and then $x_2' = x_2, \ldots, x_n' = x_n$.

3. At this point we have a basis for $V$ such that $\tau$ has matrix

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

where $A'$ is a $(n-1) \times (n-1)$ symmetric matrix. We then inductively proceed with $A'$, repeat until we're done.

### 4.2.2 Consequences of the existence of such nice bases for $K = \mathbb{C}$ or $\mathbb{R}$

**Proposition 4.2.2.** A quadratic form $q$ on n-dim vector space $V$ over $\mathbb{C}$ has the form

$$q(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_r^2$$

with respect to a suitable basis $b_1, \ldots, b_n$ for $V$ where $r =$ rank of $q$ or $\tau_q$.

*Proof.* We can write

$$q(y_1, \ldots, y_n) = \beta_1 y_1^2 + \cdots + \beta_r y_r^2 + \underbrace{\beta_{r+1} y_{r+1}^2 + \cdots + \beta_n y_n^2}_{0}.$$

And we introduce new coordinates $x_1 = \sqrt{\beta_1} y_1, \ldots, x_r = \sqrt{\beta_r} y_r$ so that we get desired in proposition. □

The matrix looks like this:

$$\begin{pmatrix} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ & 0 & & 0 \end{pmatrix}$$

**Theorem 4.2.3** (Sylvester)**.** A quadratic form $q$ on n-dim vector space $V$ over $\mathbb{R}$ has the form

$$q(x_1, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_t^2 - x_{t+1}^2 - \cdots - x_{t+u}^2$$

with respect to a suitable basis, where $t + u = r =$ rank of $q$.

The matrix looks like this:

$$\begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 0 \end{pmatrix}$$

Basically because you are in real field you cannot get rid of negatives since you don't have $i$.

**Theorem 4.2.4** (Sylvester's law of inertia)**.** Let $V$ be an $n$-dim vector space over $\mathbb{R}$, and $e_1, \ldots, e_n,\ e'_1, \ldots, e'_n$ are bases such that

$$q(x_1 e_1 + \cdots + x_n e_n) = x_1^2 + \cdots + x_t^2 - x_{t+1}^2 - \cdots - x_{t+u}^2$$

and

$$q(x_1 e'_1 + \cdots + x_n e'_n) = x_1^2 + \cdots + x_{t'}^2 - x_{t'+1}^2 - \cdots - x_{t'+u'}^2.$$

Then $t = t',\ u = u'$, i.e. numbers of $+1$ and $-1$ are uniquely determined.

*Proof.* We know that $r$ is constant, so it suffices to show that $t = t'$. We prove it by contradiction. Without loss of generality, assume $t > t'$. Define

$$V_1 := \operatorname{span}(e_1, \ldots, e_t)$$

$$V_2 := \operatorname{span}\left(e'_{t'+1}, \ldots, e'_n\right)$$

both subspaces of $V$. Now $q$ is positive on $V_1$, i.e. for $v \neq 0 \in V_1$, $q(v) > 0$, and for $v \in V_2$ $q(v) \leq 0$, so $V_1 \cap V_2 = \{0\}$. We also know that $\dim V_1 = t$ and $\dim V_2 = n - t'$. So

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2) = t + n - t' > n,$$

a contradiction since $V_1 + V_2 \subset V$. □

This justify the name "signature" of $q$ for the ordered pair $(t, u)$.

## 4.3 Euclidean vector space and orthogonal transformations

**Definition 4.3.1.** If $t = n$ (no $-1$ or 0) then we call $\tau/q$ *positive-definite*, since

$$q(v) = \tau(v, v) > 0 \qquad \forall v \neq 0 \in V.$$

**Definition 4.3.2.** In this case we call $(V, \tau)$ a *Euclidean vector space*, $\tau$ the *scalar product* of the space, and write

$$\tau(v, w) =: v \cdot w \qquad \text{or sometimes } \langle v, w \rangle.$$

**Definition 4.3.3.** We call a basis

$$e_1, \ldots, e_n : \tau(e_i, e_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

an *orthonormal basis*. (i.e. length 1 (normal) and perpendicular (ortho-) to each other.) Such bases always exist by Sylvester's theorem.

We can then write $v_E = V = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ with $E = (e_1, \ldots, e_n)$. Then $v \cdot w = v^T w = \sum_{i=1}^{n} x_i y_i$.

**Notation** (Norm/length)**.** $|v| = \sqrt{\tau(v, v)} = \sqrt{v \cdot v}$.

**Definition 4.3.4.** The angle $\varphi$ between nonzero vectors $v, w$ is defined by

$$\cos \varphi = \frac{v \cdot w}{|v||w|}$$

where $0 \leq \varphi < \pi$.

So whenever you have a Euclidean space, you can talk about lengths and angles.

**Theorem 4.3.5** (Gram–Schmidt/orthonormalisation process)**.** We start with $g_1, \ldots, g_n$, some basis of a Euclidean vector space. Then $\exists$ an orthonormal basis $f_1, \ldots, f_n$ such that $\operatorname{span}\{g_1, \ldots, g_i\} = \operatorname{span}\{f_1, \ldots, f_i\}$ for $1 \leq i \leq n$, and $f_i$'s are constructed from the $g_i$'s as followed:

1. Normalise the first vector: $f_1 = \frac{g_1}{|g_1|}$
2. Inductively, assume $f_1, \ldots, f_r$ $(1 \leq r < n)$ have been constructed. We define

$$f'_{r+1} := g_{r+1} - \sum_{i=1}^{r} (g_{r+1} \cdot f_i) f_i$$

   and normalise it: $f_{r+1} = \frac{f'_{r+1}}{|f'_{r+1}|}$.

In particular, the basis change matrix

$$\mathcal{M}(\mathrm{id}_V)_{F=(f_1,\ldots,f_n)}^{G=(g_1,\ldots,g_n)}$$

is upper triangular.

*Proof.* The equality of spans and normality of $f_{r+1}$ are by construction. We just need to check $f'_{r+1}$ is orthogonal to $f_1, \ldots, f_r$.
Compute for $1 \leq j \leq r$:

$$\begin{aligned} f'_{r+1} \cdot f_j &= g_{r+1} \cdot f_j - \sum_{i=1}^{r} (g_{r+1} \cdot f_i) \overbrace{(f_i \cdot f_j)}^{\text{nonzero only when } i=j} \\ &= g_{r+1} \cdot f_j - g_{r+1} \cdot f_j \cdot 1 = 0. \end{aligned}$$

And we know $f_{r+1}$ is not zero by linear independence of $g_i$'s. □

**Corollary 4.3.6.** If $e_1, \ldots, e_k$ is an orthonormal set of vectors in $V$, it can always be completed to an orthonormal basis $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ for $V$.

*Proof.* We can always complete $e_1, \ldots, e_k$ to a basis by first year linear algebra, then Gram–Schmidt it. □

We now just need more definitions to see the usefulness of this process.

**Definition 4.3.7.** Given $(V, \tau)$ Euclidean, a linear map $T : V \to V$ is called *orthogonal* if it preserves scalar product:

$$T(v) \cdot T(w) = v \cdot w \qquad \forall v, w \in V.$$

Suppose $A$ is the matrix of $T$ with respect to an orthonormal basis $E = (e_1, \ldots, e_n)$. We write $v_E = V,\ w_E = w$, and we have

$$T(v) \cdot T(w) = (Av)^T (Aw) = v^T A^T A w = v^T w.$$

We see that this is only possible when $A^T A = I$:

**Proposition 4.3.8.** $T$ is orthogonal if and only if its matrix $A$ with respect to some orthonormal basis satisfies

$$A^T A = A A^T = I \Leftrightarrow A^T = A^{-1}.$$

Not surprisingly we call such matrices *orthogonal*.

Observations:

- An orthogonal matrix has determinant $\pm 1$: $(\det A)^2 = \det(A^T) \det A = \det(A^T A) = \det I = 1$.
- $T : V \to V$ is orthogonal if and only if given an orthonormal basis $e_1, \ldots, e_n$ the vectors $T(e_1), \ldots, T(e_n)$ are again an orthonormal basis

**Proposition 4.3.9** ($QR$-decomposition)**.** Given an $n \times n$ matrix $A \in \mathbb{R}^{n^2}$, we can write it as a product $A = QR$ where $Q$ is orthogonal and $R$ upper triangular.

This gives us a faster algorithm solving $Ax = b$: we can just solve $Rx = Q^T b$ without having to take inverses and do eliminations.

*Proof.* First assume $A$ is invertible. Then columns of $A$ give us a basis $G = (g_1, \ldots, g_n)$ for $\mathbb{R}^n$. Then consider standard basis $E = (e_1, \ldots, e_n)$ and we can view $A$ as a basis change matrix $\mathcal{M}(\mathrm{id}_V)_G^E$. Now Gram–Schmidt $G$ to get an orthonormal basis $F$. So

$$A = \mathcal{M}(\mathrm{id}_V)_F^E \cdot \mathcal{M}(\mathrm{id}_V)_G^F.$$

Then $Q = \mathcal{M}(\mathrm{id}_V)_F^E$ and $R = \mathcal{M}(\mathrm{id}_V)_G^F$ have desired properties: $Q$ is orthogonal since the columns of $Q$ form an orthonormal set of vectors, and $\mathcal{M}(\mathrm{id}_V)_F^G$ is upper triangular by Theorem 4.3.5, so $R$, its inverse, is also upper triangular.

If $A$ is not invertible, we can write $A = A'R'$ with $A'$ invertible and $R'$ upper triangular. (Using row operations on $A$, which correspond to a sequence of multiplications by invertible matrices on the left, we can transform it into upper triangular form.) Then $A = QRR'$ by what's proved, where $RR'$ is also upper triangular. □

**Example 4.3.10.** Given

$$A = \begin{pmatrix} g_1 & g_2 & g_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & -2 \\ 2 & 0 & -1 \\ 0 & -2 & -2 \end{pmatrix}$$

is non-singular. Then we Gram–Schmidt the columns of $A$:

$$f_1 = \frac{g_1}{|g_1|} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

$$f_2' = g_2 - (g_2 \cdot f_1) f_1 = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}, \qquad f_2 = \frac{f_2'}{|f_2'|} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

$$f_3' = g_3 - (g_3 \cdot f_1) f_1 - (g_3 \cdot f_2) f_2 = \begin{pmatrix} -2 \\ -1 \\ -2 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$$

$$f_3 = \frac{f_3'}{|f_3'|} = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$$

Then

$$Q = \begin{pmatrix} -\frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & 0 & -\frac{1}{\sqrt{5}} \\ 0 & -1 & 0 \end{pmatrix} \qquad R = \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & \sqrt{5} \end{pmatrix}$$

the latter by expressing $g_i$'s with $f_i$'s.

## 4.4 Nice orthonormal bases/diagonalisation of self-adjoint operators

We saw in 4.2 that given a quadratic form $q$ on some vector space $V/K$ where $2 \neq 0$ in $K$, $\exists$ a basis $b_1, \ldots, b_n$ of $V$ such that $q(x_1, \ldots, x_n) = \alpha_1 x_1^2 + \cdots + \alpha_n x_n^2$. Question now is if $V$ is Euclidean, do we have a similarly nice orthonormal basis? Answer is yes! Moreover, if we write $q(v) = \beta(v, v)$ and choose a basis for $V$ such that $\beta$ is represented by a symmetric matrix $A$, then this answer "yes" is equivalent to that $\forall A$ real symmetric, $\exists Q$ orthogonal such that $Q^T AQ$

is diagonal: $\begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$, and since $Q$ is orthogonal, this means the above diagonal matrix is also $Q^{-1}AQ$, implying that $\alpha_i$'s are the eigenvalues of $A$.

But we'll see that the proof of this will be easier in a slightly different but equivalent language. Time for definitions.

**Definition 4.4.1.** Let $(V, \tau)$ be Euclidean. The *adjoint* linear map to $T : V \to V$ is the unique linear map $T^* : V \to V$ such that $\forall v, w \in V,\ \tau(Tv, w) = \tau(v, T^*w)$.

There's one way to justify this definition. Pick an orthonormal basis $E = (e_1, \ldots, e_n)$ and we have matrix $A$ with respect to this basis of $T$. Then we write $v, w$ in $E$ and want $(Av)^T w = v^T A^* w$. We immediately see that we can just define $A^* = A^T$ since $(Av)^T w = v^T A^T w$. But this depends very much on choice of basis and doesn't tell us much of the behind the scenes.

**Definition 4.4.2.** The *dual space* of $V$, denoted $V^*$, is the vector space of all linear forms (functionals) $l : V \to K$ on $V$.

We can convince ourselves that $\langle \cdot, \cdot \rangle$ gives an isomorphism $V \xrightarrow{\sim} V^*$. So for fixed $w_0 \in V$, $v \mapsto \langle Tv, w_0 \rangle$ gives a linear form on $V$. Then $\exists v_0$ such that this is $l_{v_0} := T^* w_0$ and $\langle T^* w_0, v \rangle = \langle v, T^* w_0 \rangle = \langle Tv, w_0 \rangle$. All this is to show that it's meaningless to talk about adjoint maps without notion of inner product.

**Definition 4.4.3.** $T$ is *self-adjoint* if $T = T^*$.

This means matrix $A$ of $T$ with respect to some orthonormal basis is symmetric.

We will prove our result about quadratic forms (symmetric matrices) in the following form.

**Theorem 4.4.4.** Given self-adjoint $T : V \to V$, $\exists$ an orthonormal basis of $V$ consisting of eigenvectors for $T$.

We first need $T$ does have an eigenvector to start with.

**Proposition 4.4.5.** Any real symmetric $n \times n$ matrix has a real eigenvalue. Moreover, all its eigenvalues are real.

*Proof.* Let $A : \mathbb{C}^n \to \mathbb{C}^n$. Then $A$ has complex eigenvalue $\lambda$ because characteristic polynomial has roots by FTA. So $Av = \lambda v$ where $v \in \mathbb{C}^n$. Now consider $\overline{v} \in \mathbb{C}^n$ and we have $\overline{A}\overline{v} = \overline{\lambda}\overline{v}$. But $A$ is real symmetric, i.e. $A = \overline{A}$. So $\overline{\lambda}$ is also an eigenvector of $A$ with eigenvector $\overline{v}$. Now consider $\overline{v}^T v$. This is positive since it's $\sum |v_i|^2$ and $v$ as an eigenvector is nonzero. So $(A\overline{v})^T v = \overline{\lambda}\overline{v}^T v = \overline{v}^T A^T v = \overline{v}^T A v = \lambda \overline{v}^T v$. Therefore $\lambda = \overline{\lambda}$. □

*Proof of Theorem 4.4.4.* By induction on $\dim V = n$.

If $n = 0$ there's nothing to do.

For $n > 0$, take an eigenvector $v \neq 0$ for $T$ with eigenvalue $\lambda$: $T(v) = \lambda v$. Let

$$\begin{aligned} W &:= \text{orthogonal complement of } v \text{ with respect to given scalar product} \\ &= \{w \in V : v \cdot w = 0\}. \end{aligned}$$

Then $\dim W = n - 1$ (seen in proof of Theorem 4.2.1). So if we can prove that $W$ is $T$-invariant, i.e. $T(W) \subset W$, we're done, since $W$ is again Euclidean and $T|_W$ is self-adjoint.

This is easy. Take $w \in W$. Then $Tv \cdot w = \lambda(v \cdot w) = 0 = v \cdot Tw$ since $T$ is self-adjoint. So $Tw \in W$. □

Right. So how do we find $Q$?

**Proposition 4.4.6.** Suppose $A$ is real symmetric $n \times n$ and we have distinct eigenvalues $\lambda, \mu$ with associated eigenvectors $v, w$. Then $v^T w = 0$.

*Proof.* Consider $(Av)^T w$. By construction it's $\lambda v^T w$. On the other hand, since $A$ is symmetric, it's $v^T A^T w = v^T A w = \mu v^T w$. But $\lambda \neq \mu$, so $v^T w$ must be zero. □

## 4.5 Geometry of quadrics in real Euclidean vector space

Consider $\mathbb{R}^n$ with the standard scalar product (or generally a Euclidean vector space $V$) and $x_1, \ldots, x_n$ coordinates. Look at zero sets of the form

$$\sum_{i=1}^{n} a_i x_i^2 + \sum_{1 \leq i \leq j \leq n} \alpha_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i x_i + \gamma = 0, \qquad (*)$$

e.g. $\mathbb{R}^4$ with $x_1, x_2, x_3, x_4$ coordinates and we look at $3x_1^2 + 5x_4^2 + 10x_1x_3 + x_2x_4 + 3x_1 + 5x_4 + 1001 = 0$. We want to classify the zero sets given like this up to isometries, up to change of coordinates of the form

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \underbrace{A}_{\text{orthogonal}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \underbrace{b}_{\text{a "translation" vector}}.$$

These zero sets are called (affine, real) quadrics, but be careful these could be counter intuitive, e.g. in $\mathbb{R}^2$, $x^2 + y^2 = -1$ is empty, or $x^2 + y^2 = 0$ is a single point.

We do this step by step. (Instead of adding apostrophes crazily, we simply call the new coordinates $x_i$ after each step.)

1. Using an appropriate isometry (with $b = 0$) we can assume that the quadratic part of $(*)$ is a sum of squares by Theorem 4.2.1.

   e.g. $x^2 + xy + y^2 +$ (linear form in $x, y$)+ constant. We can't just Babylonianly complete the square: indeed $x^2 + xy + y^2 = \left(x + \frac{1}{2}y\right)^2 + \frac{3}{4}y^2$, but this is not an orthogonal coordinate transformation.

   Note that $x^2 + xy + y^2 = (x, y) \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$, and the matrix has eigenvalues $\frac{3}{2}$, $\frac{1}{2}$ with corresponding (normal) eigenvectors $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. (We can check this also verifies Proposition 4.4.6.) Then let

   $$x := \frac{1}{\sqrt{2}}(x' + y')$$
   $$y := \frac{1}{\sqrt{2}}(x' - y')$$

   and we have

   $$\begin{aligned} x^2 + xy + y^2 &= \frac{(x')^2 + 2x'y' + (y')^2 + (x')^2 - (y')^2 + (x')^2 - 2x'y' + (y')^2}{2} \\ &= \frac{3}{2}(x')^2 + \frac{1}{2}(y')^2 \end{aligned}$$

   and again we can verify that the coefficients are precisely the eigenvalues.

2. At this point we can assume the equation is of the form

$$\sum_{i=1}^{n} a_i x_i^2 + \sum_{i=1}^{n} \beta_i x_i + \gamma = 0.$$

If one of the $\alpha_i$'s, $\alpha_r$, is nonzero, we can eliminate the summand $\beta_r x_r$ in the linear form. Particularly, we have

$$\cdots + \alpha_r x_r^2 + \cdots + \beta_r x_r + \cdots$$

and if we let

$$x_r := x_r' - \frac{\beta_r}{2\alpha_r}$$

so that the other $x_i$'s remain unchanged ($x_i' = x_i$) and $x_r' = x_r + \frac{\beta_r}{2\alpha_r}$; then we have

$$\begin{aligned}
&\cdots + \alpha_r \left( x_r' - \frac{\beta_r}{2\alpha_r} \right)^2 + \cdots + \beta_r \left( x_r' - \frac{\beta_r}{2\alpha_r} \right) + \cdots \\
&= \cdots + \alpha_r \left( x_r' \right)^2 - 2\alpha_r x_r' \frac{\beta_r}{2\alpha_r} + \frac{\alpha_r \beta_r^2}{4\alpha_r^2} + \cdots + \beta_r x_r' - \frac{\beta_r^2}{2\alpha_r} + \cdots \\
&= \cdots + \alpha_r \left( x_r' \right)^2 + \frac{\alpha_r \beta_r^2}{4\alpha_r^2} + \cdots - \frac{\beta_r^2}{2\alpha_r} + \cdots .
\end{aligned}$$

3. After possible re-indexing the $x_i$'s we reach the form

$$\alpha_1 x_1^2 + \cdots + \alpha_r x_r^2 + \underbrace{\beta_{r+1} x_{r+1} + \cdots + \beta_{r+s} x_{r+s}}_{\text{linear form}} + \gamma = 0.$$

where $\alpha_i$'s are nonzero. If the linear form is nonzero we change the coordinate with

$$\begin{array}{ccccc}
x_1 & \cdots & x_r & \dfrac{1}{\sqrt{\sum_{i=1}^{s} \beta_{r+i}^2}} & \left( \displaystyle\sum_{i=1}^{s} \beta_{r+i} x_{r+i} \right) \\
\| & & \| & & \| \\
x_1' & \cdots & x_r' & & x_{r+1}'
\end{array}$$

and $x_i'$'s form a orthonormal set. We then extend it to a basis, then Gram–Schmidt it to get an orthonormal basis $x_1', \ldots, x_{r+1}', \ldots, x_n'$. Now the linear form is just $\beta x_{r+1}'$.

4. Dividing by the coefficient of the linear form (if there is one) we can now have the form

$$\sum_{i=1}^{r} \alpha_i x_i^2 - x_{r+1} + \gamma = 0.$$

In this case we can eliminate $\gamma$ by a further translation

$$x_{r+1} := x_{r+1}' + \gamma.$$

From this we get

**Theorem 4.5.1.** After a suitable isometry/change of coordinates/rigid motions, we can write $(*)$ as either

1. $\sum_{i=1}^{r} \alpha_i x_i^2 = 0$ case where linear form is eliminated after 2. and $\gamma = 0$
2. $\sum_{i=1}^{r} \alpha_i x_i^2 = 1$ case where linear form is eliminated and $\gamma \neq 0$
3. or $\sum_{i=1}^{r} \alpha_i x_i^2 - x_{r+1} = 0$ case where linear form is not eliminated

where $1 \leq r \leq n$ and $\alpha_i$'s are nonzero.

What kind of geometrical shapes do these define? (for $n = 2,\ n = 3$)
In $\mathbb{R}^2$ we can then further classify 9 different cases - we subsume two equations in 1. 2. 3. in the same case if one arises from the other by

(A) rescaling the length scales along the axes

(B) permutation of coordinates

(C) dividing by constant $(-1)$;

also $\alpha_i > 0$:

1. $\alpha x^2 = 0$ the $y$-axis
2. $\alpha x^2 = 1$ two parallel lines $x = \pm\frac{1}{\sqrt{\alpha}}$
3. $-\alpha x^2 = 1$ empty
4. $\alpha x^2 + \beta y^2 = 0$ the single point $(0, 0)$
5. $\alpha x^2 - \beta y^2 = 0$ two lines $y = \pm\sqrt{\frac{\alpha}{\beta}}x$
6. $\alpha x^2 + \beta y^2 = 1$ ellipse
7. $\alpha x^2 - \beta y^2 = 1$ hyperbola
8. $-\alpha x^2 - \beta y^2 = 1$ empty
9. $\alpha x^2 - y = 0$ parabola

In $\mathbb{R}^3$ things get interesting. We get again the previous 9 cases, but now viewed as equations in 3 variables $x, y, z$ only not involving $z$ (cylinders). Further we have

10. $-\alpha x^2 - \beta y^2 - \gamma z^2 = 1$ empty
11. $\alpha x^2 + \beta y^2 + \gamma z^2 = 0$ the single point $(0, 0, 0)$
12. $\alpha x^2 + \beta y^2 - \gamma z^2 = 0$.

    What's this? Let's assume $\alpha = \beta = \gamma = 1$ for simplicity, then we get the general from this. If we intersect the shape with $z = c$ we get a circle $x^2 + y^2 = c^2$, and if we intersect it with $y = 0$ we get two lines $x - z = x + z = 0$ going through the origin, so this is a cone. (These

lines are sometimes called the generator of the cone.) More generally - rescaling by, say, $\begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$ - this is an elliptical cone.

13. $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$

    We get a sphere when $\alpha = \beta = \gamma = 1$, so generally we get an ellipsoid - a stretched sphere. This one doesn't have generators.

14. $\alpha x^2 + \beta y^2 - \gamma z^2 = 1$

    Again consider $x^2 + y^2 - z^2 = 1$. Again restricting $z = c$ we get circles, and when $y = 0$ we get a hyperbola going through $x = \pm 1$ when $z = 0$, so we have a hyperboloid. Since it's connected it's called hyperboloid of one sheet. The generators are not obvious but there are two families of them and every point lies on exactly one of the lines of the two families.

15. $\alpha x^2 - \beta y^2 - \gamma z^2 = 1$

    Again when $y = 0$ we have the same hyperbola. But restricting $x = c$, then if $c^2 < 1$ it's empty, and if $c^2 \geq 1$ we get circles. So it's still hyperboloid, but disconnected, so it's called hyperboloid of two sheets. (The hyperbola's other side of that of one sheet.)

16. $\alpha x^2 + \beta y^2 - z = 0$

    We get nothing below $z = 0$ and above it we have circles bounded by parabola. So in general we have what's called elliptical paraboloid.

17. $\alpha x^2 - \beta y^2 - z = 0$

    Restricting $y = 0$ we have $z = x^2$ and restricting $x = 0$ we have $z = -y^2$. So we get a (monkey (?)) saddle (hyperbolic paraboloid) and the origin is called the saddle point. Again we have 2 families of generators, meaning having 2 lines passing through each point. (Restricting $z = c$ we have two lines $x + y = x - y = z$.) See notes for details.

## 4.6 Singular value decomposition

Question: Given Euclidean vector spaces $V, W$ over $\mathbb{R}$ with inner products $\langle \cdot, \cdot \rangle_V$, $\langle \cdot, \cdot \rangle_W$ (or "$\cdot$" if there is no risk of confusion) and linear map $T : V \to W$, is there a nice form of matrix for chosen orthonormal basis in $V$ and $W$ (independently) like the following?

$$\begin{pmatrix} \gamma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \gamma_n & \\ & 0 & & 0 \end{pmatrix}$$

where $n = \operatorname{rank} T$. Answer is yes and let's formulate that.

**Theorem 4.6.1.** Given the above situation, $\exists$ orthonormal bases $e_1, \ldots, e_{\dim V}$ of $V$ and $f_1, \ldots, f_{\dim W}$ of $W$ such that

$$T(e_1) = \gamma_1 f_1 \quad \cdots \quad T(e_n) = \gamma_n f_n$$

and the remaining

$$T(e_{n+1}) = 0 \quad \cdots \quad T(e_{\dim V}) = 0$$

where $n = \text{rank } T = \dim \text{Im } T$ and $\gamma_1, \ldots, \gamma_n > 0$ are real and uniquely determined by $T$: the positive square roots of the nonzero eigenvalues of $T^*T$, where $T^*$, still called adjoint of $T$, is the unique linear map $W \to V$ such that $\forall v \in V, w \in W$, $\langle Tv, w\rangle_W = \langle v, T^*w\rangle_V$. This justifies their name, the *singular values* of $T$.

**Corollary 4.6.2** (Restatement in matrix language)**.** Given an $M \times N$ matrix $A$, $\exists$ orthogonal matrices $P, Q$ such that

$$P^T AQ = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{i.e.} \quad A = P \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} Q^T.$$

where $D = \begin{pmatrix} \gamma_1 & & & \\ & \gamma_2 & & \\ & & \ddots & \\ & & & \gamma_n \end{pmatrix}$ is a diagonal $n \times n$ matrix where $n = \text{rank } A$ and $\gamma_1 \geq \cdots \geq \gamma_n > 0$ are positive square roots of nonzero eigenvalues of $A^TA$. They are unique, meaning if we have another such decomposition

$$(P')^T AQ' = \begin{pmatrix} D' & 0 \\ 0 & 0 \end{pmatrix}$$

then the diagonal matrices of $D'$ are again the $\gamma_i$'s.

*Proof of Theorem 4.6.1.* We can define a new symmetric bilinear form on $V$:

$$v_1 * v_2 := \langle T(v_1), T(v_2)\rangle_W.$$

It may be no longer positive definite due to potential kernel, but it's at least positive semi-definite, i.e. $\forall v \in V,\ v * v \geq 0$. By Theorem 4.2.1 we can make the matrix of this bilinear form diagonal by picking a suitable orthonormal basis of $V$ $(e_1, \ldots, e_{\dim V})$. But

$$v_1 * v_2 = \langle Tv_1, Tv_2\rangle_W = \langle v_1, T^*Tv_2\rangle_V.$$

So the diagonal entries of the matrix $B$ we have are just eigenvalues

$$\alpha_1 \geq \cdots \geq \alpha_n > \alpha_{n+1} = 0 = \alpha_{n+2} = \cdots = \alpha_{\dim V}$$

(since positive semi-definite) of $T^*T$.

We still need to prove $n = \text{rank } T$. Observe that $T^*T(e_1) = \alpha_1 e_1, \ldots, T^*T(e_n) = \alpha_n e_n$. We claim that Ker $T$ = Ker $T^*T$, which follows from the observation:

$$\langle Tv, Tv\rangle_W = \langle v, T^*Tv\rangle_V,$$

so Ker $T \subset$ Ker $T^*T$ is clear. If we pick $v$ in Ker $T^*T$, then right hand side of above is zero, but inner product on $W$ is positive definite so left hand side being zero means $Tv = 0$. Now we know 2 things:

$$T(e_{n+1}) = \cdots = T(e_{\dim V}) = 0$$

since $T^*T$ kills the same vectors, and if we put $f_1' := T(e_1), \ldots, f_n' := T(e_n)$, then

$$\langle f_i', f_j'\rangle_W = e_i * e_j = \alpha_i \delta_{ij}$$

where $\delta$ is Kronecker and $\alpha_i$'s are positive. So $f_i'$'s are an orthogonal set of vectors of nonzero in $W$. Then if we normalise them to $f_i := \frac{1}{\sqrt{\alpha_i}} f_i'$ we have an orthonormal set. In particular they are linearly independent, so rank $T = n$. We can then extend it to an orthonormal basis of $W$ using Gram–Schmidt.

The only thing we haven't shown is uniqueness, but this can be done quickly. If we have a described decomposition, then $A^TA = QD^TP^TPDQ^T = QD^TDQ^{-1}$, so $D^TD = D^2$ is similar to $A^TA$, meaning they have the same eigenvalues. □

### 4.7 The complex story

Omitted, non-examinable, but recommended.

## 5 Structure theory of finitely generated abelian groups, Smith normal form

### 5.1 Definitions

**Definition 5.1.1.** An *abelian group* $G$ is a set $G$ with a map $+ : G \times G \to G$ such that

1. $\forall g_1, g_2, g_3 \in G,\ (g_1 + g_2) + g_3 = g_1 + (g_2 + g_3)$. associativity
2. $\forall g_1, g_2 \in G,\ g_1 + g_2 = g_2 + g_1$ commutativity
3. $\exists 0_G : 0_G + g = g\ \forall g \in G$ existence of neutral element

   $\Rightarrow$ $0_G$ is unique
4. $\forall g \in G,\ \exists (-g) : g + (-g) = 0_G$ existence of inverses

   $\Rightarrow$ $(-g)$ is uniquely determined

**Example 5.1.2.**

- $(\mathbb{Z}, +)$ where $+$ is the usual addition
- $(\mathbb{Z}/n, +)$ where $\mathbb{Z}/n = \{0, 1, \ldots, n-1\}$
- $(\mathbb{Q}, +)$
- More generally take $K$-vector space $V$ and $+$ the vector addition defined, then $(V, +)$ is an abelian group
- Given field $K$ then $K \backslash \{0_K\}$ with multiplication defined is an abelian group

**Definition 5.1.3.** For $n \in \mathbb{Z},\ n > 0$ and abelian group $G,\ g \in G$, define

$$n \cdot g = ng := \underbrace{g + g + \cdots + g}_{n \text{ times}}.$$

If $n < 0$,

$$ng := \underbrace{-g + (-g) + \cdots + (-g)}_{-n \text{ times}},$$

and if $n = 0$, $0 \cdot g = 0_G$.

This defines a map $\mathbb{Z} \times G \to G$.

**Definition 5.1.4.** An abelian group $G$ is *cyclic* if $\exists x \in G : G = \{nx | n \in \mathbb{Z}\}$.

**Definition 5.1.5.** A *homomorphism* between abelian groups $G_1, G_2$ is a map

$$f : G_1 \to G_2 : f(g \underbrace{+}_{G_1} h) = f(g) \underbrace{+}_{G_2} f(h) \qquad \forall g, h \in G_1.$$

$\Rightarrow f(0_{G_1}) = 0_{G_2}$ and $f(-g) = -f(g)$.

**Definition 5.1.6.** An *isomorphism* of abelian groups $G_1, G_2$ is a homomorphism $f : G_1 \to G_2$ with a two-sided inverse $g : G_2 \to G_1$ which is again a homomorphism such that $g \circ f = \mathrm{id}_{G_1},\ f \circ g = \mathrm{id}_{G_2}$.

We can check that this is equivalent to a bijective homomorphism.

**Proposition 5.1.7.** Any cyclic group $G$ is isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/n$ for some $n > 0$.

*Proof.* We know $\exists x \in G : G = \{nx | x \in \mathbb{Z}\}$. Define homomorphism

$$\begin{aligned} \phi : \mathbb{Z} &\to G \\ n &\mapsto nx \end{aligned}$$

then there are 2 cases:

1. $nx$ are all distinct, then $\phi$ is bijective hence isomorphism
2. For $n_1 \neq n_2,\ n_1 x = n_2 x$, i.e. $(n_1 - n_2)x = 0$. We pick the smallest $n \in \mathbb{Z} : nx = 0$. Then $G$ is isomorphic to $\mathbb{Z}/n$. Indeed

   $$0, x, 2x, \ldots, (n-1)x$$

   are distinct by construction, and they give all elements in $G$ which can be written as $Nx = (\alpha n + r)x = rx$ where $r \in \{0, \ldots, 1\}$. So

   $$\overline{\phi} : \mathbb{Z}/n \to G$$

   is an isomorphism.

□

**Definition 5.1.8.** $G$ abelian. The *order* of $g \in G$ is the smallest positive $n \in \mathbb{Z} : ng = 0_G$ or $\infty$ if no such $n$ exists.

**Definition 5.1.9.** If $X \subset G$ is a subset, we say $X$ *generates* $G$ if every $g \in G$ can be written as $\sum_{i=1}^{r} \alpha_i x_i$ where $\alpha_i \in \mathbb{Z},\ x_i \in X$ and $r \in \mathbb{N}$.
If $|X| < \infty$ we call $G$ *finitely generated*.

**Definition 5.1.10.** $G_1, \ldots, G_N$ abelian. The *direct sum* (or *direct product*) of the $G_i$'s is defined by

$$G_1 \times \cdots \times G_N$$

where $\times$ is Cartesian product.

From now on all groups are abelian.

## 5.2 Subgroups, cosets, quotient groups

**Definition 5.2.1.** $G$ group, $H \subseteq G$. $H$ is called a *subgroup* if it's a group with the same group operation.

**Proposition 5.2.2** (Criteria for subgroup)**.** $G$ group, $H \subseteq G$. Then $H$ is a subgroup if and only if

1. $H$ is nonempty
2. $\forall h_1, h_2 \in H,\ h_1 + h_2 \in H$
3. $\forall h \in H,\ -h \in H$

*Proof.* See notes. □

**Proposition 5.2.3** (Another one)**.** $G$ group, $H \subseteq G$. Then $H$ is a subgroup if and only if

1. $H$ is nonempty
2. $\forall h_1, h_2 \in H,\ h_1 - h_2 \in H$

*Proof.* Immediately from previous proposition. □

**Proposition 5.2.4.** If $H$ is a subgroup of $G$, then $0_H = 0_G$.

*Proof.* We know $0_H + 0_H = 0_H$ and $0_H + 0_G = 0_H$. By cancellation we have desired. □

**Example 5.2.5.**

- For any group $G$ we have 2 trivial groups $\{0_G\}$ and $G$.
- We can also take $g \in G$ and look at the cyclic group it generates: $\{ng : n \in \mathbb{Z}\} \subseteq G$.
- Even integers, denoted $2\mathbb{Z} \subseteq \mathbb{Z}$. More generally $n\mathbb{Z}$ is a subgroup.

**Definition 5.2.6.** $G$ group, $H$ subgroup, $g \in G$. We define the *coset* of $g$ by

$$H + g := \{h + g : h \in H\}.$$

**Example 5.2.7.** Let $G = \mathbb{Z},\ H = 5\mathbb{Z}$. Then $H+1 = \{5n+1 : n \in \mathbb{Z}\} = \{-4, 1, 6, \ldots\}$. Note this is also $H+(-4), H+6$, etc. We then see there are 5 cosets in total, $H, H+1, H+2, H+3, H+4$.

**Example 5.2.8.** Let $G = \mathbb{R}^2,\ H = \text{span}\{(1,2)\}$. Then coset of $(0,1)$ is $\{(x, 2x+1) : x \in \mathbb{R}\}$. This has a nice visualisation: the subgroup is a line through origin and the coset is shifted.

**Proposition 5.2.9.** $G$ group, $H$ subgroup, $g, k \in G$. Then the following are equivalent:

1. $k \in H + g$
2. $H + k = H + g$
3. $g - k \in H$

*Proof.*

- 2 implies 1: $k = 0 + k \in H + k = H + g$.

- 1 implies 2: $k = h + g$ for some $h \in H$, so for an arbitrary $h' \in H$,

$$h' + k = h' + (h + g) = (h' + h) + g \in H + g$$

  i.e. $H + k \subseteq H + g$. By symmetry of $k$ and $g$ we conclude $H + k = H + g$.

- 1 implies 3: $k = h + g$ for some $h \in H$, then $g - k = -h \in H$.
- 3 implies 1: $g - k = h$ for some $h \in H$, then $k = -h + g \in H + g$.

□

**Corollary 5.2.10.** Given $g_1, g_2$, either $H + g_1 = H + g_2$ or $H + g_1 \cap H + g_2 = \varnothing$.

*Proof.* Suppose not disjoint, i.e. $k \in H + g_1 \cap H + g_2$. By equivalence of 1 and 2, $H + g_1 = H + k = H + g_2$. □

**Corollary 5.2.11.** Cosets of $H$ partition $G$.

**Proposition 5.2.12.** All cosets have the same cardinality $|H|$.

*Proof.* Consider a map $H \to H + g,\ h \mapsto h + g$. Clearly it's surjective. Now suppose $h_1, h_2 \in H$ such that $h_1 + g = h_2 + g$. By cancellation $h_1 = h_2$, so it's injective and therefore bijective. □

**Corollary 5.2.13** (Lagrange theorem)**.** If $G$ is finite, $|G| = |H|$(number of cosets of $H$)

Number of cosets of is then called index of $H$.

**Proposition 5.2.14.** $|g|$ divides $|G|$.

*Proof.* Consider the cyclic subgroup $H$ generated by $G$ and by previous corollary we have desired. □

**Proposition 5.2.15.** If $|G| = p$ where $p$ is prime, then $G$ is cyclic.

**Notation.** $G/H = \{\text{cosets of } H\}$.

**Definition 5.2.16.** $G$ group, $A, B \subseteq G$. We define $A + B := \{a + b, a \in A, b \in B\}$.

$\Rightarrow\ A + B \subseteq G$

**Lemma 5.2.17.** $H$ subgroup, then $(H + g) + (H + k) = H + (g + k)$.

*Proof.* $h_1 + g + h_2 + k = (h_1 + h_2) + (g + k)$. □

**Theorem 5.2.18** (Quotient group)**.** $G/H$ together with addition defined above is an abelian group.

*Proof.* Closure by previous lemma, associativity and commutativity by those of $G$, $0_{G/H} = H$ and inverse of $H + g$ is $H + (-g)$. □

## 5.3 First isomorphism theorem

**Definition 5.3.1.** Given homomorphism $\phi : G \to H$, *kernel* is defined by Ker $\phi = \{g \in G : \phi(g) = 0_H\}$.

**Proposition 5.3.2.** $\phi$ is injective if and only if Ker $\phi = \{0_G\}$.

*Proof.* If injective then $\phi(0_G) = 0_H$.
Conversely suppose Ker $\phi = \{0_G\}$. Then $\phi(g_1) - \phi(g_2) = \phi(g_1 - g_2) = 0_H$, so $g_1 - g_2 = 0_G$, so injective. □

**Theorem 5.3.3.** Given $\phi : G \to H$, Ker $\phi$ is a subgroup of $G$ and Im $\phi$ is a subgroup of $H$.

**Theorem 5.3.4.** Define $\phi : G \to G/H,\ g \mapsto H + g$. This is a surjective homomorphism with Ker $\phi = H$.

**Proposition 5.3.5.** $\phi : G \to H$ a homomorphism with kernel $K$, $A \subseteq G$ a subgroup. Then the following are equivalent:

1. $A \subseteq K$
2. $\exists \overline{\phi} : G/A \to H,\ \overline{\phi}(A + g) = \phi(g)$.

*Proof.*

- 1 implies 2: suppose $A + g_1 = A + g_2$. By previous proposition $g_1 - g_2 \in A \subseteq K$. So $\phi(g_1 - g_2) = 0_H \Rightarrow \phi(g_1) = \phi(g_2)$. So the $\overline{\phi}$ in 2 is well-defined, i.e. it's consistent. To show homomorphism, note that

$$\begin{aligned}\overline{\phi}((A + g) + (A + h)) = \overline{\phi}(A + g + h) = \phi(g + h) = \phi(g) + \phi(h)\\ = \overline{\phi}(A + g) + \overline{\phi}(A + h)\end{aligned}$$

- not 1 implies not 2: suppose $\exists a \in A : \phi(a) \neq 0$. We can write

$$\overline{\phi}(A + a) = \overline{\phi}(A) = \phi(a) \neq 0_H$$

but $A$ is identity of $G/A$, so $\overline{\phi}$ is ill-defined.

□

**Example 5.3.6.** Let $\phi : \mathbb{Z} \to \mathbb{Z}/3$ be $n \mapsto n \mod 3$. So $K = 3\mathbb{Z}$, and let $A = 6\mathbb{Z} \subseteq K$. So $A + 1 = \{\ldots, -5, 1, 7, 13, \ldots\}$, then $\phi(a) = 1\ \forall a \in A + 1$. But if we take $A' = 5\mathbb{Z} \not\subseteq K$, $A' + 1 = \{\ldots, -4, 1, 6, 11, \ldots\}$ then $\phi(a') = 0, 1, 2\ \forall a' \in A' + 1$, conflicting answers which means making sense of $\overline{\phi}$ is impossible.

**Theorem 5.3.7** (First isomorphism theorem)**.** $\phi : G \to H$ with kernel $K$. Then $\overline{\phi} : G/K \to H$ gives an isomorphism of $G/K \cong$ Im $\phi$.

*Proof.* Existence of the homomorphism stated is a immediate corollary of the previous proposition. It follows that Im $\overline{\phi}$ = Im $\phi$, so $\overline{\phi} : G/K \to$ Im $\phi$ is surjective. Now suppose $(K + g) \in$ Ker $\overline{\phi}$. So $\overline{\phi}(K + g) = \phi(g) = 0_H$, i.e. $g \in K$. This implies $K + g = K$ by Proposition 5.2.9, which is identity of $G/K$. So Ker $\overline{\phi} = \{0_{G/K}\}$, therefore $\overline{\phi}$ is injective, hence bijective. □

## 5.4 Finitely generated free abelian groups

**Definition 5.4.1.** We call an abelian group *free* of rank $n$ if it is isomorphic to $\mathbb{Z}^n$ (with componentwise addition). ($n \in \mathbb{N}_{>0}$)

**Definition 5.4.2.** Let $G$ be an abelian group. We call elements $g_1, \ldots, g_n \in G$

- *(integrally) linearly independent* if whenever there is a relation

$$\alpha_1 g_1 + \alpha_2 g_2 + \cdot + \alpha_n g_n = 0_G$$

  where $\alpha_i \in \mathbb{Z}$, we must have $\alpha_1 = \cdots = \alpha_n = 0$;
- *(integrally) span/generate* $G$ if every element in $G$ can be written as a $\mathbb{Z}$-linear combination of $g_i$'s.
- *an (integral) basis* if they have both above properties.

**Example 5.4.3.** Let $G = \mathbb{Z}^n$ and $g_1, \ldots, g_r \in G,\ r \in \mathbb{N}_{>0}$ where $g_i = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ with $a_i \in \mathbb{Z}$. We can view these as elements of $\mathbb{Q}^n$, which is a vector space, since $\mathbb{Z}^n \subset \mathbb{Q}^n$. We claim $g_1, \ldots, g_r$ are integrally linearly independent if and only if they are linearly independent considered as vectors in $\mathbb{Q}^n$.

*Proof.* $\Leftarrow$: If we write

$$\alpha_1 g_1 + \cdots + \alpha_r g_r = 0 \quad \in \quad \mathbb{Z}^n$$

and when $g_1, \ldots, g_r$ are $\mathbb{Q}$-linearly independent, we get $\alpha_1 = \cdots = \alpha_r = 0$ by definition since $\mathbb{Q}$ and $\mathbb{Z}$ share the same zero.

$\Rightarrow$: We prove this by proving if $g_1, \ldots, g_r$ are linearly dependent in $\mathbb{Q}^n$ then they are integrally linearly dependent too. If we write

$$\frac{p_1}{q_1} g_1 + \frac{p_2}{q_2} g_2 + \cdots + \frac{p_r}{q_r} g_r = 0,$$

a nontrivial linear dependency relation in $\mathbb{Q}^n$ where $p_i \in \mathbb{Z},\ q_i \in \mathbb{N}_{>0}$. Multiply by all the $q_i$'s we get a nontrivial linear dependency relation with integral coefficients.

□

We further claim that if $g_1, \ldots, g_r$ span $\mathbb{Z}^n$ integrally then $g_1, \ldots, g_r$ span $\mathbb{Q}^n$ as a vector space.

*Proof.* Let $v = \begin{pmatrix} \frac{a_1}{b_1} \\ \vdots \\ \frac{a_n}{b_n} \end{pmatrix} \in \mathbb{Q}^n$ where $a_i \in \mathbb{Z}$ and $b_i \in \mathbb{N}_{>0}$. Then $\prod b_i v \in \mathbb{Z}^n$ is a linear combination of $g_1, \ldots, g_r$. So if we divide by $\prod b_i$ again we have $v$ being a $\mathbb{Q}$-linear combination of $g_1, \ldots, g_r$. □

Converse is not true. Counterexample: $\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix}$ span $\mathbb{Q}^2$ but does not span $\mathbb{Z}^2$ since you only get even tuples.

**Remark.** Suppose $g_1, \ldots, g_r$ is an integral basis of $G \simeq \mathbb{Z}^n$, then $r = n$. i.e. all integral basis have $n$ elements. Indeed, by preceding claims, $g_1, \ldots, g_r$ is a basis for $\mathbb{Q}^n$, so $r = n$.

**Theorem 5.4.4.** Suppose $x_1, \ldots, x_n$ are an integral basis of $G = \mathbb{Z}^n$ and $y_1, \ldots, y_n \in G$. We can form an $n \times n$ matrix $P$ whose columns contain the basis expansion of $y_1, \ldots, y_n$ with respect to the basis $x_1, \ldots, x_n$. Then the following are equivalent:

1. $y_1, \ldots, y_n$ are also an integral basis
2. $P$ is invertible with $P^{-1} \in \mathbb{Z}^{n \times n}$
3. $\det P = \pm 1$

*Proof.*

- $1 \Rightarrow 2$: Let $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n)$. Then $P$ is just change of basis matrix $\mathcal{M}(\mathrm{id})_Y^X$, so $\mathcal{M}(\mathrm{id})_X^Y \cdot \mathcal{M}(\mathrm{id})_Y^X = I_n$. Since $Y$ is an integral basis, $\mathcal{M}(\mathrm{id})_X^Y = P^{-1} \in \mathbb{Z}^{n \times n}$.
- $2 \Rightarrow 3$: $PP^{-1} = I_n$, so $\det P \det P^{-1} = 1$. But both determinants are integers, so we have 3.
- $3 \Rightarrow 2$: Since $P^{-1} = \frac{1}{\det P} \mathrm{adj}\ P$ where adj $P$ is polynomial in entries of $P$ and $\det P = \pm 1$, so entries of $P^{-1}$ are integers.
- $2 \Rightarrow 1$: We can express $x_j$'s as integral linear combination of $y_i$'s, i.e. $Y$ spans $\mathbb{Z}^n$ integrally. Also $Y$ is $\mathbb{Q}$-linearly independent so it's integrally linearly independent.

□

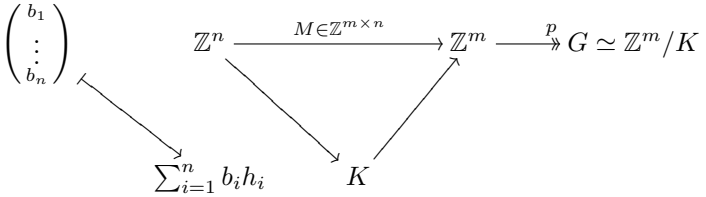
## 5.5 (Unimodular) Smith normal form for integer matrices, structure of finitely generated abelian groups

### 5.5.1 Motivation and preparation

$$K = \mathrm{Ker}\ p \longrightarrow \mathbb{Z}^m \xrightarrow{p \text{ surjective}} G$$

$$\text{abelian subgroup of } \mathbb{Z}^n \qquad \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \longmapsto a_1 g_1 + \cdots + a_m g_m$$

where $G$ is a finitely generated abelian group with $g_1, \ldots, g_m$ generators, Ker $p$ will be proved to be finitely generated. But then we can expand this by iterating:




where $K$ is finitely generated by $h_1, \ldots, h_n$ and is $\mathrm{Im}(M) \subseteq \mathbb{Z}^m$.

This gives us a free presentation of $G$,

$$\mathbb{Z}^n \xrightarrow{M} \mathbb{Z}^m \twoheadrightarrow \mathbb{Z}^m/\mathrm{Im}(M) \simeq G$$

which says $G$ is generated by $g_1, \ldots, g_m$ and modulo relations given by columns of $M$.

**Example 5.5.1.**

$$\mathbb{Z}^2 \xrightarrow{\left(\begin{smallmatrix} 2 & -1 \\ 3 & 2 \\ 1 & 4 \end{smallmatrix}\right)} \mathbb{Z}^3 \twoheadrightarrow \mathbb{Z}^m/\mathrm{Im}(M) \simeq G$$

which says $G$ is generated by 3 elements $g_1, g_2, g_3$ subject to relations

$$\begin{aligned} 2g_1 + 3g_2 + g_3 &= 0 \\ -g_1 + 2g_2 + 4g_3 &= 0 \end{aligned}$$

in $G$.

Then suppose we want to understand the isomorphism type of $G$, what we do is to bring $M$ to a particularly "easy" normal form using integrally invertible row and column operations.
Over $\mathbb{Q}$ we can transform $M$ into Smith normal form

$$\begin{pmatrix} 1 & & & 0 & \\ & 1 & & & \\ & & \ddots & & 0 \\ 0 & & & 1 & \\ & & 0 & & \end{pmatrix}$$

where there are $r = \operatorname{rank} M$ of 1's. But something like dividing by 2 is not considered as invertible over $\mathbb{Z}$. We regulate the following to be allowed:

- Add an integer multiple of some row/column to another
- Interchange two rows/columns
- Multiply a row/column by an $-1$

These are sometimes called unimodular row and column operations. These correspond to multiplication of $M$ on the left/right by matrices in $GL_m(\mathbb{Z})/GL_n(\mathbb{Z})$, $m \times m/n \times n$ matrices with integer entries and inverse also with integer.

**Example 5.5.2.** $A = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \in GL_2(\mathbb{Z})$ since $\det A = 1$, but $B = \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix} \notin GL_2(\mathbb{Z})$ since $\det B = -10$.

### 5.5.2 Smith normal form

**Theorem 5.5.3** (Unimodular Smith normal form)**.** Given $M \in \mathbb{Z}^{m\times n}$, we can reduce $M$ to the form by unimodular row and column operations:

$$\left(\begin{array}{cccc|c} d_1 & & & 0 & \\ & d_2 & & & \\ & & \ddots & & 0 \\ 0 & & & d_r & \\ \hline & & 0 & & \end{array}\right)$$

where $d_1, \dots, d_r \in \mathbb{Z}_{>0}$, $d_i | d_{i+1}$ $\forall i = 1, \dots, r-1$ and $r = \operatorname{rank} M$. Furthermore the form is unique.

This allows us to determine the isomorphism type of a general finitely generated abelian group. Consider again the presentation and suppose $M'$ is the unimodular Smith norm form of $M$. Then we have

$$\begin{array}{ccccc} \mathbb{Z}^n & \xrightarrow{M} & \mathbb{Z}^m & \twoheadrightarrow & \mathbb{Z}^m/\mathrm{Im}(M) \simeq G \\ \downarrow{\scriptstyle B\in GL_n(\mathbb{Z})} & & \downarrow{\scriptstyle A\in GL_m(\mathbb{Z})} & & \uparrow{\scriptstyle \text{isomorphism induced by } A} \\ \mathbb{Z}^n & \xrightarrow[M']{} & \mathbb{Z}^m & \twoheadrightarrow & \mathbb{Z}^m/\mathrm{Im}(M') \end{array}$$

where $M' = AMB^{-1}$. So now we just need to understand $\mathbb{Z}^m/\mathrm{Im}(M')$ which is isomorphic to $G$. We write the generators $g_1, \dots, g_m$ modulo the only relations

$$\begin{aligned} d_1 g_1 &= 0 \\ d_2 g_2 &= 0 \\ &\vdots \\ d_r g_r &= 0, \end{aligned}$$

no other relations. So we can see $\mathbb{Z}^m/\mathrm{Im}(M')$ as direct sum of cyclic groups $\mathbb{Z}/d_i\mathbb{Z}$'s with $\mathbb{Z}^{n-r}$, a free group generated by remaining $g_{r+1}, \dots, g_m$ where $d_i$ is just order of the cyclic group.

*Proof of Theorem 5.5.3.* Consider the set $S$ of all integer $m \times n$ matrices that we can attain from $M$ by unimodular row/column operations and $M$ is assumed to be nonzero. In this set, pick $\widetilde{M}$ with the property that one of its entries has minimum absolute value among all entries of all matrices in $S$ and is nonzero.

$$\widetilde{M} = \left(\begin{array}{c|c|c} & y & \\ \hline y & x & \\ \hline & & \end{array}\right)$$

We claim that other entries in the same row and column are all divisible by $x$. Indeed, if $y = qx + r$ and $0 < |r| < |x|$, then we can just add $-q$ times the row/column of $x$ to the row/column of $y$

and get a contradiction, so $r = 0$. Therefore we can make all of them zero and have

$$\widetilde{\widetilde{M}} = \left(\begin{array}{c|c|c} & 0 & \\ & \vdots & \\ \hline 0 \quad \cdots & x & \cdots \quad 0 \\ \hline & \vdots & \\ & 0 & \end{array}\right)$$

and then by reordering we put $x$ in $1,1$ spot:

$$\left(\begin{array}{c|ccc} x & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & M \in \mathbb{Z}^{m-1 \times n-1} & \\ 0 & & & \end{array}\right)$$

and we proceed by induction on $m+n$ to reach the normal form with $d_1, \ldots, d_r > 0$. Now suppose $d_1$ would not divide $d_2$, but again by division with remainder and row/column operations we can come up with an entry with smaller absolute value than $d_1 = x$, a contradiction. By induction $d_i$ divides $d_{i+1}$. □

How to actually compute this form?

Observe that given $M \in \mathbb{Z}^{m \times n}$, $d_1$ in the Smith normal form is just the gcd of all the entries. To see this we want to show that this gcd is invariant under unimodular row/column operations (then if it's gcd of all $d_i$'s then it's gcd of original entries). Indeed, if we add an integer multiple of one row/column to another and get $\widetilde{M}$ from $M$, then an integer divides all entries of $M$ if and only if it divides all entries of $\widetilde{M}$. So here's the algorithm.

1. Compute gcd of $M \in \mathbb{Z}^{m \times n}$. If the gcd is actually $\pm$one of entries, we can just make other entries sharing the row and column 0 and move it to 1,1.

   **Example 5.5.4.**

   $$\begin{pmatrix} 8 & 6 \\ 4 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 4 & -2 \\ 8 & 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 4 & 2 \\ 8 & -6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 4 \\ -6 & 8 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 0 \\ -6 & 20 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 0 \\ 0 & 20 \end{pmatrix}$$

2. If none of the entries are equal to gcd, we try to reduce the size of absolute value minimum entry (using division with remainder and row/column operations) until gcd appears.

   $$\left(\begin{array}{c|c|c} & y & \\ \hline y & x & \\ \hline & & \end{array}\right)$$

   2 cases: either $x$ divides all the $y$ sharing the row and column, or not. If one of $y$ is not divisible by $x$, then just do row/column operations and go back to step 1.

   **Example 5.5.5.**

   $$\begin{pmatrix} 7 & 13 & 4 \\ 6 & 2 & 5 \\ 10 & 20 & 102 \end{pmatrix}$$

If all of them are divisible, then make all $y$ zero. There will be another nonzero entry $z$ outside the cross band that is not divisible by $x$ by assumption that $x$ is not gcd. This gives us a submatrix

$$\begin{pmatrix} z & 0 \\ 0 & x \end{pmatrix}$$

and since $z = qx + r$ where $0 < |r| < |x|$, we can replace $z$ by $r$ by row/column operations (add row of $x$ to row of $z$ and add $-q$ times column of $x$ to column of $z$). We do this until gcd appears and go back to step 1.

**Example 5.5.6.**

$$\begin{pmatrix} -18 & -18 & -18 & 90 \\ 54 & 12 & 45 & 48 \\ 9 & -6 & 6 & 63 \\ 18 & 6 & 15 & 12 \end{pmatrix}$$

gcd is 3. To make it appear we stare at the matrix and see that we can subtract column 3 from 1.

$$\begin{pmatrix} 0 & -18 & -18 & 90 \\ 9 & 12 & 45 & 48 \\ 3 & -6 & -6 & 63 \\ 3 & 6 & 15 & 12 \end{pmatrix}$$

and by row/column operations we arrive at

$$\left(\begin{array}{c|ccc} 3 & 0 & 0 & 0 \\ \hline 0 & -6 & 0 & 12 \\ 0 & -12 & -9 & 51 \\ 0 & -18 & -18 & 90 \end{array}\right)$$

and gcd of submatrix is again 3. So we make column 1 positive and add column 2 to 1 to get

$$\begin{pmatrix} 6 & 0 & 12 \\ 3 & -9 & 51 \\ 0 & -18 & 90 \end{pmatrix}$$

then we arrive at

$$\left(\begin{array}{c|cc} 3 & 0 & 0 \\ \hline 0 & 18 & -90 \\ 0 & -18 & 90 \end{array}\right)$$

and proceed similarly we conclude at

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

### 5.5.3 Structure of finitely generated abelian groups

Recall with Smith normal forms, $\mathbb{Z}/\mathrm{Im}(M')$ is just

$$\mathbb{Z}/d_1\mathbb{Z} \oplus \mathbb{Z}/d_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_r\mathbb{Z} \oplus \mathbb{Z}^{n-r}$$

and $\mathbb{Z}/1\mathbb{Z}$ is just the trivial group so can be omitted (unless it's the only member). ($K$, a subgroup of finitely generated group is also finitely generated is remained to be proved.) We can then write the main theorem of this section.

**Theorem 5.5.7** (Structure theorem for finitely generated abelian groups)**.** Each finitely generated abelian group $G$ is isomorphic to a group of the form

$$\mathbb{Z}/a_1\mathbb{Z} \oplus \mathbb{Z}/a_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_s\mathbb{Z} \oplus \mathbb{Z}^f$$

where $a_i \in \mathbb{Z},\ a_i \geq 2$ and $a_i | a_{i+1}\ \forall i$; and $f \in \mathbb{N}$; and $s, f, a_i$'s are uniquely determined by $G$.

**Example 5.5.8.** Consider a finitely generated abelian groups of order 36. Then immediately $f = 0$ since otherwise it would be infinite. Also $a_1 \cdots a_s = 36$ where $a_i$'s are of properties described in theorem. Trivially we can have $a_1 = 36$. Also $a_1 = 2,\ a_2 = 18;\ a_1 = 3,\ a_2 = 12$ or $a_1 = a_2 = 6$. These are different because, say, $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/18\mathbb{Z}$ has elements of order 18 while others don't.

But if we have a group of order 16, then it could be $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z}$ or $\mathbb{Z}/4\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z}$. Then we can count number of elements of order 2 in each isomorphism and find that one has 7 while the other only has 3.

**Lemma 5.5.9.** Given a product of abelian groups

$$G = G_1 \times G_2 \times \cdots \times G_N$$

and an element

$$g = (g_1, g_2, \ldots, g_N)$$

of finite order $o(g)$, then $o(g) = \text{lcm } \{o(g_i)\}$.

*Proof.* Let $l := \text{lcm } \{o(g_i)\}$. Then $g^l = \left(g_1^l, \ldots, g_N^f\right) = 0_G$. So $o(g)|l$. Now suppose $g^r = (g_1^r, \ldots, g_N^r) = 0_G$. Then $o(g_i)|r\ \forall i$, so $\text{lcm } \{o(g_i)\} := l|r$. Put $r = o(g)$ we complete the proof. □

**Remark.** Many naturally occurring abelian groups are not finitely generated, e.g. $(\mathbb{Q}, +)$.

Now we'll prove the unproved.

**Proposition 5.5.10.** A subgroup $H$ of a finitely generated abelian group $G$ is finitely generated.

*Proof.* Suppose $G$ is minimally generated by nonzero elements $x_1, \ldots, x_n$. We prove the desired by induction on $n$. If $n = 0$ then $G$ is trivial so it's trivial. Then we look at the subgroup $G'$ of $G$ generated by $x_1, \ldots, x_{n-1}$. We have 2 alternatives:

1. $H \subseteq G'$, then we're okay by induction.
2. $H \not\subseteq G'$, then $\exists g' + tx_n \in H$ where $g' \in G',\ t \in \mathbb{Z}_{>0}$, i.e. there is an element in $H$ which is not generated by $x_1, \ldots, x_{n-1}$ and we do need $x_n$ to write it. We pick such element with minimal $t$. Consider $H \cap G' \subset G'$ which can be generated by $y_1, \ldots, y_n$ and denote $g' + tx_n := y_{n+1}$. We claim $H$ is generated by $y_1, \ldots, y_n, y_{n+1}$. Consider a general $\widetilde{g} + sx_n \in H$ where $\widetilde{g} \in G',\ s \in \mathbb{Z}$. $s$ must be multiple of $t$ by minimality of latter and division with remainder:

**Example 5.5.11.** Suppose $G' = \langle x_1, x_2, x_3 \rangle$ and $y_{n+1} = 2x_1 + 3x_2 + 5x_3$. But if we have $4x_1 + 101x_2 + 7x_3$ then $2x_1 + 98x_2 + 2x_3 \in H$ but $2 < 5$, a contradiction;

so $\widetilde{g} + sx_n - Ny_{n+1} \in G' \cap H$.

□

**Remark.** 1. Consider $R$-modules $M$ (which satisfy same axioms with $K$-vector space, except scalars are taken from $R$, a commutative ring) then abelian groups are just $Z$-modules. So why is the proof above nontrivial? Is a submodule of any finitely generated $R$-module finitely generated? No! Consider $R = K[x_1, x_2, \ldots]$, a polynomial ring over field $K$ in infinitely many variables, then a $R$-submodule generated by $x_1, x_2, \ldots$ is not finitely generated. We exploit the speciality, namely division with remainder, of $\mathbb{Z}$.

2. If we have 2 isomorphic finitely generated abelian groups $G_1 \simeq \mathbb{Z}/a_1\mathbb{Z} \oplus \mathbb{Z}/a_r\mathbb{Z} \oplus \mathbb{Z}^e$ and $G_2 \simeq \mathbb{Z}/b_1\mathbb{Z} \oplus \mathbb{Z}/b_s\mathbb{Z} \oplus \mathbb{Z}^f$ with $a_i, b_j \geq 2$ and $a_i | a_{i+1},\ b_j | b_{j+1}\ \forall i, j$. Then $e = f,\ r = s$ and $a_i = b_i$, i.e. unique determined by the group.

   *Sketch of proof.*

$$\begin{array}{ccc} G_1 & \xrightarrow{f} & G_2 \\ \cup & \simeq & \cup \\ T_1 & & T_2 \end{array}$$

   where $T_1, T_2$ are "torsion subgroups", i.e. subgroups formed by all elements of finite order. In this case $T_1 = \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_r\mathbb{Z}$ and $T_2 = \mathbb{Z}/b_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/b_s\mathbb{Z}$. Clearly $f$ restricts to an isomorphism $f : T_1 \xrightarrow{\sim} T_2$ and also induces an isomorphism $\mathbb{Z}^e = G_1/T_1 \xrightarrow{\sim} G_2/T_2 = \mathbb{Z}^f$. Take an element of maximum order in $T_1$, say $x$, then under $f$ it's mapped to the element of maximum order in $T_2$, say $y$, i.e. we must have $a_r = b_s$. We then see that $f$ induces an isomorphism

$$\mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_{r-1}\mathbb{Z} = T_1/\langle x \rangle \xrightarrow{\sim} T_2/\langle y \rangle = Z/b_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/b_{s-1}\ ,$$

   then we proceed by induction. □

3. There is also a similar "unimodular" Smith normal form where you replace $\mathbb{Z}$ with polynomials in one variables $K[x]$ (which also has division with remainder). What interesting $\mathbb{C}[x]$ modules are there where structure could be illuminated by the Smith normal form for polynomials?

   Consider a finite dimensional $\mathbb{C}$-vector space $V$ with endomorphism $T$. We define $p(x)v$ where $v \in V$ and $p(x) \in \mathbb{C}[x]$ by

$$\left(a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 \mathrm{id}\right)(v).$$

   This $\rightsquigarrow$ the Jordan canonical form. Essentially. We proved the same theorem twice (assuming a sufficiently conceptual point of view) in 2 different ways.