The little book around Linear algebra

Arun Ram   email: aram@unimelb.edu.au

August 29, 2026

Abstract

These are notes towards *The Little Book around Linear Algebra*. I am very grateful to Ruirui Yao, Avi Tendler and Eric Wang for their very helpful careful proofreading and insightful remarks.

Key words — Matrices, Row reduction, Vector spaces, Inner products

Contents

AMS Subject Classifications: Primary 05E??; Secondary 33D??.

1 Matrices and operations

1.1 Matrices

Let $\mathbb{F}$ be a field. Let $s, t \in \mathbb{Z}_{>0}$.

- An $t \times s$ matrix with entries in $\mathbb{F}$ is a table of elements of $\mathbb{F}$ with t rows and s columns. More precisely, an $t \times s$ matrix with entries in $\mathbb{F}$ is a function
$$A\colon \{1, \ldots, t\} \times \{1, \ldots, s\} \longrightarrow \mathbb{F}.$$
- A *column vector of length* t is an $t \times 1$ matrix.
- A *row vector of length* s is an $1 \times s$ matrix.
- The (i, j) *entry of a matrix* A is the element $A(i, j)$ in row i and column j of A.
$$A = \begin{pmatrix} A(1,1) & A(1,2) & \cdots & A(1,s) \\ A(2,1) & A(2,2) & \cdots & A(2,s) \\ \vdots & & & \vdots \\ A(t,1) & A(t,2) & \cdots & A(t,s) \end{pmatrix}$$

Let

$M_{t\times s}(\mathbb{F})$ denote the set of $t \times s$ matrices with entries in $\mathbb{F}$.

1.2 Addition and scalar multiplication

Let $t, s \in \mathbb{Z}_{>0}$.

- The *sum* of $t \times s$ matrices A and B is the $t \times s$ matrix $A + B$ given by
$$(A + B)(i, j) = A(i, j) + B(i, j), \qquad \text{for } i \in \{1, \ldots, t\} \text{ and } j \in \{1, \ldots, s\}.$$
- The *scalar multiplication* of an element $c \in \mathbb{F}$ with an $t \times s$ matrix A is the $t \times s$ matrix $c \cdot A$ given by
$$(c \cdot A)(i, j) = c \cdot A(i, j), \qquad \text{for } i \in \{1, \ldots, t\} \text{ and } j \in \{1, \ldots, s\}.$$

The *zero matrix* is the $t \times s$ matrix $0 \in M_{t\times s}(\mathbb{F})$ given by
$$0(i, j) = 0, \qquad \text{for } i \in \{1, \ldots, t\} \text{ and } j \in \{1, \ldots, s\}.$$

The *negative* of a matrix $A \in M_{t\times s}(\mathbb{F})$ is the matrix $-A \in M_{t\times s}(\mathbb{F})$ given by
$$(-A)(i, j) = -A(i, j), \qquad \text{for } i \in \{1, \ldots, t\} \text{ and } j \in \{1, \ldots, s\}.$$

Proposition 1.1. *Let $t, s \in \mathbb{Z}_{>0}$ and let $M_{t\times s}(\mathbb{F})$ be the set of $t \times s$ matrices with entries in $\mathbb{F}$.*

(a) If $A, B, C \in M_{t\times s}(\mathbb{F})$ then $A + (B + C) = (A + B) + C$.

(b) If $A, B \in M_{t\times s}(\mathbb{F})$ then $A + B = B + A$.

(c) If $A \in M_{t\times s}(\mathbb{F})$ then $0 + A = A$ and $A + 0 = A$.

(d) If $A \in M_{t\times s}(\mathbb{F})$ then $(-A) + A = 0$ and $A + (-A) = 0$.

(e) If $A \in M_{t\times s}(\mathbb{F})$ and $c_1, c_2 \in \mathbb{F}$ then $c_1 \cdot (c_2 \cdot A) = (c_1c_2) \cdot A$.

(f) If $A \in M_{t\times s}(\mathbb{F})$ and $1 \in \mathbb{F}$ is the identity in $\mathbb{F}$ then $1 \cdot A = A$.

1.3 Multiplication

- The *product* of an $t \times s$ matrix A and an $s \times p$ matrix B is the $t \times p$ matrix AB given by

$$\begin{aligned}(AB)(i,k) &= \sum_{j=1}^{n} A(i,j)B(j,k) \\ &= A(i,1)B(1,k) + A(i,2)B(2,k) + \cdots + A(i,n)B(n,k),\end{aligned}$$

  for $i \in \{1,\ldots,t\}$ and $k \in \{1,\ldots,p\}$.

The *Kronecker delta* is given by

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Let $n \in \mathbb{Z}_{>0}$. The *identity matrix* is the $n \times n$ matrix $1 \in M_{n\times n}(\mathbb{F})$ given by

$$1(i,j) = \delta_{ij}, \qquad \text{for } i \in \{1,\ldots,n\} \text{ and } j \in \{1,\ldots,n\}.$$

Proposition 1.2. *Let $n \in \mathbb{Z}_{>0}$ and let $M_n(\mathbb{F})$ be the set of $n \times n$ matrices in $\mathbb{F}$.*

(a) If $A, B, C \in M_n(\mathbb{F})$ then $A + (B + C) = (A + B) + C$.

(b) If $A, B \in M_n(\mathbb{F})$ then $A + B = B + A$.

(c) If $A \in M_n(\mathbb{F})$ then $0 + A = A$ and $A + 0 = A$.

(d) If $A \in M_n(\mathbb{F})$ then $(-A) + A = 0$ and $A + (-A) = 0$.

(e) If $A, B, C \in M_n(\mathbb{F})$ then $A(BC) = (AB)C$.

(f) If $A, B, C \in M_n(\mathbb{F})$ then $(A + B)C = AC + BC$ and $C(A + B) = CA + CB$.

(g) If $A \in M_n(\mathbb{F})$ then $1A = A$ and $A1 = A$.

1.4 The favorite basis of $M_{t\times s}(\mathbb{F})$

Let $t, s \in \mathbb{Z}_{>0}$. For $k \in \{1,\ldots,t\}$ and $\ell \in \{1,\ldots,s\}$ let $E_{k\ell} \in M_{t\times s}(\mathbb{F})$ be the matrix given by

$$E_{k\ell}(i,j) = \begin{cases} 1, & \text{if } i = k \text{ and } j = \ell, \\ 0, & \text{otherwise,} \end{cases}$$

so that $E_{k\ell}$ has a 1 in the (k,ℓ) entry and all other entries 0.

The following Proposition says that every matrix can be written as a unique combination (using scalar multiplication and addittion) of the E_{ij}.

Proposition 1.3. *Let $t, s, p \in \mathbb{Z}_{>0}$ and let $A \in M_{t\times s}(\mathbb{F})$. Then there exist unique $c_{ij} \in \mathbb{F}$, for $i \in \{1,\ldots,t\}$ and $j \in \{1,\ldots,s\}$ such that*

$$A = \sum_{i=1}^{t} \sum_{j=1}^{s} c_{ij} E_{ij}.$$

1.5 Matrix transpose

The *transpose* of an $t \times s$ matrix A is the $s \times t$ matrix A^T given by

$$A^T(i,j) = A(j,i), \qquad \text{for } i \in \{1,\ldots,s\} \text{ and } j \in \{1,\ldots,t\}.$$

Proposition 1.4. *Let $t, s, n \in \mathbb{Z}_{>0}$, let $M_{t\times s}(\mathbb{F})$ be the set of $t \times s$ matrices with entries in $\mathbb{F}$, and let $M_n(\mathbb{F})$ be the set of $n \times n$ matrices in $\mathbb{F}$.*

(a) If $A, B \in M_{t\times s}(\mathbb{F})$ then $(A+B)^T = A^T + B^T$,

(b) If $A \in M_{t\times s}(\mathbb{F})$ and $c \in \mathbb{F}$ then $(c \cdot A)^T = c \cdot A^T$,

(c) If $A, B \in M_n(\mathbb{F})$ then $(AB)^T = B^T A^T$.

(d) If $A \in M_n(\mathbb{F})$ then $(A^T)^T = A$.

1.6 Symmetric, skew-symmetric, Hermitian, orthogonal and unitary matrices

For $x + yi \in \mathbb{C}$ define $\overline{x+yi} = x - yi$. For $A \in M_{t\times s}(\mathbb{C})$ define $\overline{A}^T \in M_{s\times t}(\mathbb{C})$ by

$$\overline{A}^T(i,j) = \overline{A(j,i)}, \qquad \text{for } i \in \{1,\ldots,s\} \text{ and } j \in \{1,\ldots,t\}.$$

Let $n \in \mathbb{Z}_{>0}$.

- A *symmetric matrix* is $A \in M_{n\times n}(\mathbb{F})$ such that $A = A^T$.
- A *skew-symmetric matrix* is $A \in M_{n\times n}(\mathbb{F})$ such that $A = -A^T$.
- A *Hermitian matrix* is $A \in M_{n\times n}(\mathbb{C})$ such that $A = \overline{A}^T$.
- An *orthogonal matrix* is $A \in M_{n\times n}(\mathbb{F})$ such that $AA^T = 1$.
- A *unitary matrix* is $A \in M_{n\times n}(\mathbb{C})$ such that $A\overline{A}^T = 1$.

1.7 Traces

Let $A \in M_{n\times n}(\mathbb{Q})$. The *trace of A* is

$$\mathrm{tr}(A) = A_{11} + \cdots + A_{nn} \quad \text{where} \quad A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}$$

Proposition 1.5 (Properties of trace)**.** *Let $A, B \in M_{n\times n}(\mathbb{Q})$ and let $c \in \mathbb{Q}$. Then*

$$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B), \qquad \mathrm{tr}(cA) = c\,\mathrm{tr}(A) \qquad \textit{and} \qquad \mathrm{tr}(AB) = \mathrm{tr}(BA).$$

1.8 Determinants

Let $n \in \mathbb{Z}_{>0}$. Let E_{ij} be the $n\times n$ matrix with 1 in the (i,j) entry and 0 elsewhere. For $i \in \{1,\ldots,n-1\}$ and $c \in \mathbb{Q}$ define

$$s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{ii}.$$

For $i \in \{1,\ldots,n\}$ and $d \in \mathbb{Q}$ with $d \neq 0$ define

$$h_i(d) = 1 + (d-1)E_{ii},$$

For $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $c \in \mathbb{Q}$ define

$$x_{ij}(c) = 1 + cE_{ij},$$

For $r \in \{1, \ldots, n\}$ define

$$1_r = E_{11} + \cdots + E_{rr}.$$

The *deteminant* is the function $\det\colon M_{n\times n}(\mathbb{Q}) \to \mathbb{Q}$ determined by

$$\text{if } A, B \in M_{n\times n}(\mathbb{Q}) \text{ then} \quad \det(AB) = \det(A)\det(B),$$

and if $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $k \in \{1, \ldots, n-1\}$ and $c, d \in \mathbb{Q}$ with $d \neq 0$ then

$$\det(x_{ij}(c)) = 1, \quad \det(h_i(d)) = d, \quad \det(s_k(c)) = -1.$$

Theorem 1.6 (Determinant and trace are conjugacy invariants)**.** *Let $n \in \mathbb{Z}_{>0}$. Let $A \in M_{n\times n}(\mathbb{Q})$ and let $P \in GL_n(\mathbb{Q})$ so that P is an invertible $n \times n$ matrix. Then*

$$\det(PAP^{-1}) = \det(A) \qquad \textit{and} \qquad \mathrm{tr}(PAP^{-1}) = \mathrm{tr}(A).$$

1.8.1 The permutation formula for the determinant

Let $n \in \mathbb{Z}_{>0}$. A *permutation* is an $n \times n$ matrix such that each row and each column contain exactly one 1 and all other entries are 0. If $n = 2$ then the permtutations are the matrices in the set

$$S_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

If $n = 3$ then the permutations are the matrices in the set

$$S_3 = \left\{ \begin{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{matrix} \right\}.$$

Let w be a permutation and let $w(i)$ denote the row number of the 1 in column j of w. Let

$$\mathrm{Inv}(w) = \{(i, j) \mid i, j \in \{1, \ldots, n\} \text{ and } i < j \text{ and } w(i) > w(j)\} \quad \text{and let} \quad \ell(w) = \#\mathrm{Inv}(w).$$

The permutation for the determinant of a matrix A is

$$\det(A) = \sum_{w \in S_n} (-1)^{\ell(w)} A(1, w(1)) A(2, w(2)) \cdots A(n, w(n)).$$

If $n = 2$ the permutation formula says

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \underset{\begin{pmatrix} \color{red}{a_{11}} & a_{12} \\ a_{21} & \color{red}{a_{22}} \end{pmatrix}}{(-1)^0 a_{11} a_{12}} + \underset{\begin{pmatrix} a_{11} & \color{red}{a_{12}} \\ \color{red}{a_{21}} & a_{22} \end{pmatrix}}{(-1)^1 a_{12} a_{21}} = a_{11}a_{22} - a_{12}a_{21}.$$

If $n=3$ the permutation formula says

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{pmatrix} = \begin{array}{ccccc} (-1)^0 a_{11}a_{22}a_{33} & + & (-1)^2 a_{12}a_{23}a_{31} & + & (-1)^2 a_{13}a_{21}a_{32} \\ +(-1)^1 a_{12}a_{21}a_{33} & + & (-1)^1 a_{11}a_{23}a_{32} & + & (-1)^3 a_{13}a_{22}a_{31}. \end{array}$$

corersponding to

$$\begin{array}{ccc}
\begin{pmatrix} \color{red}{a_{11}} & a_{12} & a_{13} \\ a_{21} & \color{red}{a_{22}} & a_{23} \\ a_{31} & a_{32} & \color{red}{a_{33}}\end{pmatrix} &
\begin{pmatrix} a_{11} & \color{red}{a_{12}} & a_{13} \\ a_{21} & a_{22} & \color{red}{a_{23}} \\ \color{red}{a_{31}} & a_{32} & a_{33}\end{pmatrix} &
\begin{pmatrix} a_{11} & a_{12} & \color{red}{a_{13}} \\ \color{red}{a_{21}} & a_{22} & a_{23} \\ a_{31} & \color{red}{a_{32}} & a_{33}\end{pmatrix} \\
\begin{pmatrix} a_{11} & \color{red}{a_{12}} & a_{13} \\ \color{red}{a_{21}} & a_{22} & a_{23} \\ a_{31} & a_{32} & \color{red}{a_{33}}\end{pmatrix} &
\begin{pmatrix} \color{red}{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & \color{red}{a_{23}} \\ a_{31} & \color{red}{a_{32}} & a_{33}\end{pmatrix} &
\begin{pmatrix} a_{11} & a_{12} & \color{red}{a_{13}} \\ a_{21} & \color{red}{a_{22}} & a_{23} \\ \color{red}{a_{31}} & a_{32} & a_{33}\end{pmatrix}
\end{array}$$

So

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{pmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}-a_{13}a_{22}a_{31}.$$

1.8.2 Cofactor formulas for the determinant

Let $A\in M_{n\times n}(\mathbb{Q})$ and let $i,j\in\{1,\ldots,n\}$.

Let $A^{(i,j)}$ be the matrix A with the ith row removed and the jth column removed.

The *(i,j)-cofactor of A* is

$$C_{ij}=(-1)^{i+j}\det(A^{(i,j)}).$$

Theorem 1.7. ***(cofactor expansion)*** *Let $A\in M_{n\times n}(\mathbb{Q})$ and let $i,j\in\{1,\ldots,n\}$. Then*

$$\det(A)=A_{i1}C_{i1}+A_{i2}C_{i2}+\cdots+A_{in}C_{in}$$ *cofactor expansion across the ith row*

and

$$\det(A)=A_{1j}C_{1j}+A_{2j}C_{2j}+\cdots+A_{nj}C_{nj}.$$ *cofactor expansion down the jth column*

2 Invertible matrices and factoring

2.1 Invertible matrices

A matrix $P \in M_n(\mathbb{F})$ is *invertible* if there exists a matrix $P^{-1} \in M_n(\mathbb{F})$ such that

$$PP^{-1} = 1 \qquad \text{and} \qquad P^{-1}P = 1.$$

The *general linear group* is

$$GL_n(\mathbb{F}) = \{P \in M_n(\mathbb{F}) \mid P \text{ is invertible}\}.$$

Proposition 2.1. *If $P, Q \in GL_n(\mathbb{F})$ then*

$$(PQ)^{-1} = Q^{-1}P^{-1}.$$

2.2 Factoring matrices

Let $t, s \in \mathbb{Z}_{>0}$ and let $E_{ij} \in M_{t\times s}(\mathbb{F})$ denote the matrix with 1 in entry (i,j) and 0 elsewhere, For $r \in \{1, \ldots, \min(s,t)\}$ let

$$1_r \in M_{t\times s}(\mathbb{F}) \qquad \text{be given by} \qquad 1_r = E_{11} + \cdots + E_{rr}. \tag{1rdefn}$$

For example,

$$1_4 = \begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix} \qquad \text{in } M_{7\times 15}(\mathbb{Q}).$$

The factoring algorithm, or row reduction algorithm, of Section 2.5 proves the following factorization theorem. There are more detailed statements of the output of the factoring algorithm, but the following is what is needed to compute solutions of systems of linear equations.

Theorem 2.2. *Let $s, t \in \mathbb{Z}_{>0}$ and let E_{ij} denote the $t \times s$ matrix with 1 in the (i,j) entry and 0 elsewhere.*
Let $A \in M_{t\times s}(\mathbb{F})$. Then there exist

$$r \in \{1, \ldots, \min(s,t)\}, \qquad P \in GL_t(\mathbb{F}), \qquad Q \in GL_s(\mathbb{F})$$

$$\textit{such that} \qquad A = P1_rQ, \qquad \textit{where} \quad 1_r = E_{11} + \cdots + E_{rr}.$$

2.3 Reduced row echelon form

Let $t, s \in \mathbb{Z}_{>0}$. A matrix $R \in M_{t\times s}(\mathbb{F})$ is in *reduced row echelon form*, or *left orbit representative*, if there exist $r \in \{0, 1, \ldots, \min(s, t)\}$ and $1 \le f_1 < \cdots < f_r \le s$ such that

(a) If $i \in \{1, \ldots, r\}$ then

$$\begin{array}{rr} & R(1, f_i) = 0, \\ & R(2, f_i) = 0, \\ & \vdots \\ & R(i-1, f_i) = 0, \\ R(i,1) = 0, \quad R(i,2) = 0, \quad \ldots \quad R(i, f_i - 1) = 0, & R(i, f_i) = 1, \end{array}$$

(b) If $i \in \{r+1, \ldots, t\}$ then $R(i, 1) = 0, R(i, 2) = 0, \ldots, R(i, s) = 0$.

For example,

$$R = \begin{pmatrix} 0 & 0 & 1 & 8 & 3 & 0 & 9 & 1 & 3 & 4 & 0 & 2 & 6 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 & 5 & 2 & 2 & 8 & 0 & 4 & 1 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 6 & 3 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

is a matrix in $M_{7\times 15}(\mathbb{Q})$ which is in reduced row echelon form. In this example, $t = 7$, $s = 15$, $r = 4$ and $f_1 = 3$, $f_2 = 6$, $f_3 = 11$, $f_4 = 14$.

Theorem 2.3. *Let $t, s \in \mathbb{Z}_{>0}$ and let $A \in M_{t\times s}(\mathbb{F})$. Then there exist an invertible matrix $P \in GL_t(\mathbb{F})$ and a unique matrix $R \in M_{t\times s}(\mathbb{F})$ in reduced row echelon form such that*

$$A = PR.$$

2.4 Invertible generator matrices

In order to factor matrices, it is important to first specify what the "irreducible" or "prime" factors are. Then a general matrix can be factored into these special factors. Use the terminology "generator matrices" for the favorite "prime" factors to use. The *generator matrices* are the *diagonal generators*, the *root matrices* and the *row reducers* defined as follows.

Let $t \in \mathbb{Z}_{>0}$ and let E_{ij} be the matrix in $M_{t\times t}(\mathbb{F})$ with 1 in the (i,j) entry and all other entries 0. Let

$$\mathbb{F}^\times = \{d \in \mathbb{F} \mid d \neq 0\}.$$

The *diagonal generators* are $h_i(d) = 1 + (d-1)E_{ii}$ for $i \in \{1,\ldots,t\}$ and $d \in \mathbb{F}^\times$.

$$h_i(d) = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& d &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \qquad \text{with} \qquad h_i(d)^{-1} = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& d^{-1} &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}$$

The *root matrices* are $x_{ij}(c) = 1 + cE_{ij}$ for $c \in \mathbb{F}$ and $i,j \in \{1,\ldots,t\}$ with $i < j$.

$$x_{ij}(c) = \begin{pmatrix} 1 &&&& \\ & \ddots && c & \\ && \ddots && \\ &&& \ddots & \\ &&&& 1 \end{pmatrix} \qquad \text{with} \qquad x_{ij}(c)^{-1} = \begin{pmatrix} 1 &&&& \\ & \ddots && -c & \\ && \ddots && \\ &&& \ddots & \\ &&&& 1 \end{pmatrix}$$

The *row reducers* are $s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{i,i}$ for $i \in \{1,\ldots,t-1\}$ and $c \in \mathbb{F}$.

$$s_i(c) = \begin{pmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 &&&&& \\ &&& c & 1 &&& \\ &&& 1 & 0 &&& \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{pmatrix} \qquad \text{with} \qquad s_i(c)^{-1} = \begin{pmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 &&&&& \\ &&& 0 & 1 &&& \\ &&& 1 & -c &&& \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{pmatrix}$$

The factoring algorithm presented in the next section explicitly factors A as

$$\begin{aligned} A &= ((\text{product of } s_i(c)) \cdot (\text{product of } h_i(c)) \cdot (\text{product of } x_{ij}(c)) \\ &\qquad \cdot 1_r \cdot (\text{product of } s_i(0)) \cdot (\text{product of } x_{ij}(c)) \\ &= P1_rQ = PR. \end{aligned}$$

2.4.1 Example: An upper triangular matrix in terms of root generators

Every upper triangular matrix with 1s on the diagonal is a product of $x_{ij}(c)$s. For example,

$$\begin{pmatrix} 1 & 10 & 9 & 7 & 4 \\ 0 & 1 & 8 & 6 & 3 \\ 0 & 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = x_{45}(1)x_{35}(2)x_{25}(3)x_{15}(4)x_{34}(5)x_{24}(6)x_{14}(7)x_{23}(8)x_{13}(9)x_{12}(10).$$

2.4.2 Example: A northwest upper triangular matrix in terms of row reducers

$$s_4(1)s_3(2)s_2(3)s_1(4)\cdot s_4(5)s_3(6)s_2(7)\cdot s_4(8)s_3(9)\cdot s_4(10) = \begin{pmatrix} 4 & 7 & 9 & 10 & 1 \\ 3 & 6 & 8 & 1 & 0 \\ 2 & 5 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

2.4.3 Example: A diagonal matrix in terms of diagonal generators

$$h_1(15)h_2(-3)h_3(76)h_4(-19)h_5(2) = \begin{pmatrix} 15 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 76 & 0 & 0 \\ 0 & 0 & 0 & -19 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

2.5 The factoring algorithm

Let $t, s \in \mathbb{Z}_{>0}$ and set A be a $t\times s$ matrix. The first-nonzero sequence of A is $(f_1, \ldots, f_t)$ given by

$$f_i = \begin{cases} \infty, & \text{if all entries in row } i \text{ of } A \text{ are } 0, \\ j, & \text{if the first nonzero entry in row } i \text{ of } A \text{ is in column } j. \end{cases}$$

The matrix A is in *echelon form* if $(f_1, \ldots, f_t)$, the first-nonzero sequence of A, is strictly increasing with all ∞ at the end.

Step 1: Row reducers. Assume A is not in echelon form. Let $(f_1, \ldots, f_t)$ be the first-nonzero sequence of A and let $f_0 = 0$. Let

$$m = \min\{f_k \mid f_{k-1} \ge f_k\} \qquad \text{and let} \qquad i = \max\{j \mid f_j = m\}.$$

Let

$$B = s_i\Big(\frac{A(i-1, f_i)}{A(i, f_i)}\Big)^{-1} A \qquad \text{so that} \quad A = s_i\Big(\frac{A(i-1, f_i)}{A(i, f_i)}\Big) B.$$

Repeat step 1 until the output is in echelon form.

Step 2: Diagonal generators Assume A is in echelon form. Let $(f_1, \ldots, f_t)$ be the first-nonzero sequence of A. Let

$$B = h_n(d_n^{-1})\cdots h_1(d_1^{-1})A, \qquad \text{where} \quad d_i = \begin{cases} A(i, f_i), & \text{if } f_i \ne \infty, \\ 1, & \text{if } f_i = \infty. \end{cases}$$

so that $A = h_1(d_1)\cdots h_n(d_n)B$.

Step 3: Root matrices. Assume A is in echelon form and the first nonzero entry in each nonzero row is equal to 1. Let $(f_1,\ldots,f_t)$ be the first-nonzero sequence of A. Assume that there exists $i\in\{1,\ldots,n\}$ and $j\in\{1,\ldots,f_i-1\}$ such that $A(i,j)\neq 0$. Let

$$i = \max\{k \mid f_k\neq\infty \text{ and there exists } j\in\{1,\ldots,f_k-1\} \text{ with } A(j,f_k)\neq 0\}$$

and let

$$B = x_{1,f_i}(-A(1,f_i))\cdots x_{f_i-1,f_i}(-A(f_i-1,f_i))A$$

so that $A = x_{f_i-1,f_i}(A(f_i-1,f_i))\cdots x_{1,f_i}(A(1,f_i))B$. Repeat step 3 until the output is a left GL-representative.

Step 4: Root matrices on the right. Assume R is a left GL-representative and that R has a nonzero entry that is not the first nonzero entry in its row. Let $(f_1,\ldots,f_t)$ be the first-nonzero sequence of R. Let

$$i = \min\{k \mid f_k\neq\infty \text{ and there exists } j\in\{f_k+1,\ldots,s\} \text{ such that } R(i,j)\neq 0\}.$$

Let

$$B = Rx_{i,s}(-R(i,s))x_{i,s-1}(-R(i,s-1))\cdots x_{i,f_i+1}(-R(i,f_i+1))$$

so that $R = Bx_{i,f_i+1}(-R(i,f_i+1))\cdots x_{i,s}(-R(i,s))$. Repeat Step 4 until the output is such that every nonzero row has only one nonzero entry.

Step 5: Row reducers on the right. Assume R is a left GL-representative and that every nonzero row has only one nonzero entry and that nonzero entry is equal to 1. Let $(f_1,\ldots,f_s)$ be the first-nonzero sequence of R. Let $m=\min\{i \mid f_i>i\}$. Let

$$B = Rs_{f_m-1}(0)\cdots s_{m+1}(0)s_m(0) \qquad \text{so that} \quad R = Bs_m(0)s_{m+1}(0)\cdots s_{f_m-1}(0).$$

Repeat Step 5 until $f_i=i$ for every i such that $f_i\neq\infty$.

Final result. This process explicitly factors A as

$$\begin{aligned} A &= ((\text{product of } s_i(c))\cdot(\text{product of } h_i(c))\cdot(\text{product of } x_{ij}(c)) \\ &\qquad \cdot 1_r\cdot(\text{product of } s_i(0))\cdot(\text{product of } x_{ij}(c)) \\ &= P1_rQ = PR. \end{aligned}$$

2.5.1 A factoring algorithm example – Step 1

This is an example of five iterations of Step 1 of the factoring algorithm on the matrix

$$A = \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ -9 & -108 & -\frac{233}{2} & -114 & \frac{293}{2} & -167 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix}.$$

Let

$$s_1(c) = \begin{pmatrix} c & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad s_2(c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad s_3(c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} A &= \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ -9 & -108 & -\frac{233}{2} & -114 & \frac{293}{2} & -167 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \\
&= s_2(-3) \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{2} & 12 & \frac{23}{2} & -23 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \\
&= s_2(-3)s_1(-1) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & \frac{1}{2} & 12 & \frac{23}{2} & -23 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \\
&= s_2(-3)s_1(-1)s_3(2) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \end{pmatrix} \\
&= s_2(-3)s_1(-1)s_3(2)s_2(0) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & -1 & -36 \end{pmatrix} \\
&= s_2(-3)s_1(-1)s_3(2)s_2(0)s_3(1) \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

2.5.2 A factoring algorithm example – Steps 2 and 3

This is an example of Step 2 and two iterations of Step 3 of the factoring algorithm on the matrix

$$B = \begin{pmatrix} 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \\ 0 & 0 & 0 & 0 & -1 & -36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let

$$h_1(d) = \begin{pmatrix} d & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad h_2(d) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad h_3(d) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and let

$$x_{23}(c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & c & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad x_{13}(c) = \begin{pmatrix} 1 & 0 & c & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad x_{12}(c) = \begin{pmatrix} 1 & c & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned} B &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \begin{pmatrix} 1 & 12 & 13 & 14 & 15 & 16 \\ 0 & 0 & 1 & 24 & 25 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &= h_1(3)h_2(\tfrac{1}{4})h_3(-1) \cdot x_{23}(25)x_{13}(15)x_{12}(13) \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

2.5.3 A factoring algorithm example – Steps 4 and 5

This is an example of Steps 4 and 5 of the factoring algorithm on the matrix

$$R = \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} R &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} x_{12}(12)x_{14}(14)x_{16}(16)x_{24}(24)x_{26}(26)x_{36}(36) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} s_2(0)x_{12}(12)x_{14}(14)x_{16}(16)x_{24}(24)x_{26}(26)x_{36}(36) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} s_3(0)s_4(0)s_2(0)x_{12}(12)x_{14}(14)x_{16}(16)x_{24}(24)x_{26}(26)x_{36}(36) \\ &= 1_3 \cdot s_3(0)s_4(0)s_2(0)x_{12}(12)x_{14}(14)x_{16}(16)x_{24}(24)x_{26}(26)x_{36}(36) = 1_3 Q, \end{aligned}$$

where

$$Q = \begin{pmatrix} 1 & 12 & 0 & 14 & 0 & 16 \\ 0 & 0 & 1 & 24 & 0 & 26 \\ 0 & 0 & 0 & 0 & 1 & 36 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = s_3(0)s_4(0)s_2(0)x_{12}(12)x_{14}(14)x_{16}(16)x_{24}(24)x_{26}(26)x_{36}(36).$$

2.5.4 A factoring algorithm example – The output of the factoring algorithm

The full output of the factoring algorithm on the matrix

$$A = \begin{pmatrix} -3 & -36 & -39 & -42 & -46 & -84 \\ -9 & -108 & -\frac{233}{2} & -114 & \frac{293}{2} & -167 \\ 3 & 36 & 39 & 42 & 45 & 48 \\ 0 & 0 & \frac{1}{4} & 6 & \frac{25}{4} & \frac{13}{2} \end{pmatrix} \qquad \text{is} \qquad A = P1_3Q,$$

where

$$P = \begin{matrix} s_2(-3)s_1(-1)s_3(2)s_2(0)s_3(1) \\ \cdot h_1(3)h_2(\tfrac{1}{4})h_3(-1)\cdot x_{23}(25)x_{13}(15)x_{12}(13) \end{matrix} \qquad \text{in } GL_4(\mathbb{F})$$

and

$$Q = \begin{matrix} x_{36}(36)x_{26}(26)x_{16}(16) \\ \cdot x_{25}(24)x_{15}(14)x_{24}(12) \\ \cdot s_3s_4s_2. \end{matrix} \qquad \text{in } GL_6(\mathbb{F}).$$

3 Kernels, images and solutions of linear systems

3.1 Kernels and images

Let $\mathbb{F}$ be a field. For $s \in \mathbb{Z}_{>0}$ let

$$\mathbb{F}^s = M_{s\times 1}(\mathbb{F}).$$

A *subspace of* $\mathbb{Q}^s$ is a subset $W \subseteq \mathbb{Q}^s$ such that

(a) $0 \in W$,

(b) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(c) If $w \in W$ and $c \in \mathbb{Q}$ then $cw \in W$.

If $r \in \mathbb{Z}_{>0}$ and $v_1, \ldots, v_r \in \mathbb{F}^s$ then

$$\mathbb{F}\text{-span}\{v_1, \ldots, v_r\} = \{c_1v_1 + \cdots + c_rv_r \mid c_1, \ldots, c_r \in \mathbb{F}\} \tag{3.1}$$

is the smallest subspace of $\mathbb{F}^s$ containing $\{v_1, \ldots, v_r\}$.

Let $A \in M_{t\times s}(\mathbb{F})$. Define

$$\ker(A) = \{v \in \mathbb{F}^s \mid Av = 0\} \qquad \text{and} \qquad \mathrm{im}(A) = \{Av \mid v \in \mathbb{F}^s\}. \tag{3.2}$$

Proposition 3.1. *Let* $A \in M_{t\times s}(\mathbb{F})$.

(a) $\ker(A)$ *is a subspace of* $\mathbb{F}^s$.

(b) $\mathrm{im}(A)$ *is a subspace of* $\mathbb{F}^t$.

Let V be an $\mathbb{F}$-subspace of $\mathbb{F}^s$. A $\mathbb{F}$-*basis of* V is a collection B of column vectors such that every element of V can be written as a unique combination (using scalar multiplication and addition) of the column vectors in B. In other words, a $\mathbb{F}$-*basis of* V is a set $\{b_1, \ldots, b_d\} \subseteq V$ such that

$$\text{if } v \in V \text{ then} \quad v = a_1b_1 + \cdots + a_db_d, \qquad \text{for unique } a_1, \ldots, a_d \in \mathbb{F}.$$

The dimension of V is the number of elements in a basis of V. In other words,

$$\text{if } \{b_1, \ldots, b_d\} \text{ is a } \mathbb{F}\text{-basis of } V \quad \text{then} \quad \dim(V) = d.$$

The following proposition determines $\ker(A)$ and $\mathrm{im}(A)$ in terms of its normal form decomposition $A = P1_rQ$.

Proposition 3.2. *Let* $t, s \in \mathbb{Z}_{>0}$. *Let* $r \in \{1, \ldots, \min(s,t)\}$ *and set*

$$1_r = E_{11} + \cdots + E_{rr} \quad \text{in } M_{t\times s}(\mathbb{F}).$$

Let $A \in M_{t\times s}(\mathbb{F})$. *Assume* $r \in \{1, \ldots, \min(s,t)\}$ *and*

$$P \in GL_t(\mathbb{F}) \qquad \text{and} \qquad Q \in GL_s(\mathbb{F}) \qquad \text{are such that} \qquad A = P1_rQ.$$

Then

the last $s - r$ *columns of* Q^{-1} *are an* $\mathbb{F}$*-basis of* $\ker(A)$, *and*
the first r *columns of* P *are an* $\mathbb{F}$*-basis of* $\mathrm{im}(A)$.

The following corollary is sometimes termed 'the rank-nullity theorem'.

Corollary 3.3. *Let* $A \in M_{t\times s}(\mathbb{F})$. *Then*

$$\dim(\ker(A)) + \dim(\mathrm{im}(A)) = (\textit{number of columns of } A).$$

3.2 Solutions of systems of linear equations

Using matrix multiplication the system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s &= b_2, \\ &\vdots \\ a_{t1}x_1 + a_{m2}x_2 + \cdots + a_{ts}x_s &= b_t, \end{aligned}$$

is written in the form

$$Ax = b, \qquad \text{where} \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & & & \vdots \\ a_{t1} & a_{t2} & \cdots & a_{ts} \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_t \end{pmatrix}.$$

The *set of solutions of* $Ax = b$ is

$$\mathrm{Sol}(Ax = b) = \{x \in \mathbb{F}^s \mid Ax = b\}.$$

The following proposition says that if A is square and invertible then

$$x = A^{-1}Ax = A^{-1}b \qquad \text{is the unique solution to the system of equations } Ax = b,$$

so that $\mathrm{Sol}(Ax = b)$ contains only one element.

Proposition 3.4. *Let $A \in M_{t\times s}(\mathbb{F})$ and $b \in \mathbb{F}^t$. If $t = s$ and $A \in GL_t(\mathbb{F})$ then*

$$\mathrm{Sol}(Ax = b) = \{A^{-1}b\}.$$

The following proposition says that if $\mathrm{Sol}(Ax = b)$ is nonempty then $\mathrm{Sol}(Ax = b)$ is the same size as $\ker(A)$.

Proposition 3.5. *Let $A \in M_{t\times s}(\mathbb{F})$ and $b \in \mathbb{F}^t$ and assume $\mathrm{Sol}(Ax = b) \neq \emptyset$. Let $p \in \mathrm{Sol}(Ax = b)$. Then*

$$\mathrm{Sol}(Ax = b) = p + \ker(A).$$

The following proposition determines $\mathrm{Sol}(Ax = b)$ explicitly.

Proposition 3.6. *Let $A \in M_{t\times s}(\mathbb{F})$ and $b \in \mathbb{F}^t$. Assume $r \in \{1, \ldots, \min(s,t)\}$ and $P \in GL_t(\mathbb{F})$ and $Q \in GL_s(\mathbb{F})$ are such that*

$$A = P1_rQ, \qquad \textit{and let} \qquad Q^{-1} = \begin{pmatrix} | & & | \\ q_1 & \cdots & q_s \\ | & & | \end{pmatrix}$$

so that $q_1, \ldots, q_s$ are the columns of Q^{-1}. Then, with $\mathbb{F}$-span as defined in (3.1),

$$\mathrm{Sol}(Ax = b) = \begin{cases} Q^{-1}\begin{pmatrix} (P^{-1}b)_1 \\ \vdots \\ (P^{-1}b)_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \mathbb{F}\text{-span}\left\{ \begin{pmatrix} | \\ q_{r+1} \\ | \end{pmatrix}, \ldots, \begin{pmatrix} | \\ q_s \\ | \end{pmatrix} \right\}, & \begin{array}{l} \textit{if the last } t-r \textit{ entries of } P^{-1}b \\ \textit{are equal to } 0, \end{array} \\ \emptyset, & \textit{otherwise.} \end{cases}$$

4 Eigenvalues, eigenvectors, diagonalization and orthogonality

4.1 Eigenvalues and eigenvectors

Let $A \in M_n(\mathbb{F})$.

- An *eigenvalue of* A is an element $\lambda \in \mathbb{F}$ such that $\ker(A-\lambda) \neq 0$.

Let $A \in M_n(\mathbb{F})$ and $\lambda \in \mathbb{F}$.

- An *eigenvector of* A *of eigenvalue* λ is a nonzero element of $\ker(A-\lambda)$.

4.2 Diagonalization with eigenvectors

Vectors $p_1, \ldots, p_n \in \mathbb{F}^n$ are *linearly independent* if the following condition holds:

$$\text{if } v \in \mathbb{F}^n \text{ then there exists unique } c_1, \ldots, c_n \in \mathbb{F} \text{ such that } v = c_1p_1 + \cdots c_np_n.$$

In other words, $p_1, \ldots, p_n \in \mathbb{F}^n$ are linearly independent if every vector $v \in \mathbb{F}^n$ can be written in a unique way as a linear combination of $p_1, \ldots, p_n$.

Theorem 4.1. *Let $n \in \mathbb{Z}_{>0}$ and let $A \in M_n(\mathbb{F})$.*
The matrix A has n linearly independent eigenvectors $p_1, \ldots, p_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if $A = PDP^{-1}$ where,

$$P = \begin{pmatrix} | & & | \\ p_1 & \cdots & p_n \\ | & & | \end{pmatrix} \qquad \textit{and} \qquad D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $p_1, \ldots, p_n$ are the columns of P and D is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$.

4.3 Diagonalization with orthonormal eigenvectors

Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. The standard symmetric inner product on $\mathbb{F}^n$ is the function $\langle, \rangle \colon \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by

$$\langle u, v\rangle = v^tu = u_1v_1 + \cdots + u_nv_n, \qquad \text{for} \quad \begin{matrix} u = (u_1, \ldots, u_n)^T \in \mathbb{F}^n, \text{ and} \\ v = (v_1, \ldots, v_n)^T \in \mathbb{F}^n. \end{matrix} \tag{4.1}$$

The standard Hermitian inner product on $\mathbb{F}^n$ is the function $\langle, \rangle \colon \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by

$$\langle u, v\rangle = \overline{v}^tu = u_1\overline{v_1} + \cdots + u_n\overline{v_n}, \qquad \text{for} \quad \begin{matrix} u = (u_1, \ldots, u_n)^T \in \mathbb{F}^n, \text{ and} \\ v = (v_1, \ldots, v_n)^T \in \mathbb{F}^n. \end{matrix} \tag{4.2}$$

A set $\{u_1, \ldots, u_n\}$ of vectors in $\mathbb{F}^n$ is *orthonormal* if the set $\{u_1, \ldots, u_n\}$ satisfies

$$\text{if } i, j \in \{1, \ldots, n\} \qquad \text{then} \qquad \langle u_i, u_j\rangle = \delta_{ij}.$$

Theorem 4.2. *Let $n \in \mathbb{Z}_{>0}$ and let $A \in M_n(\mathbb{F})$.*

(a) Assume that the inner product on $\mathbb{F}^n$ is given by (4.1). The matrix A has n orthonormal eigenvectors $q_1, \ldots, q_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if $A = QDQ^{-1}$ and $QQ^t = 1$ where,

$$Q = \begin{pmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{pmatrix} \qquad \textit{and} \qquad D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $u_1, \ldots, u_n$ are the columns of Q and D is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$.

(b) *Assume that the inner product on* $\mathbb{F}^n$ *is given by* (4.2)*. The matrix* A *has* n *orthonormal eigenvectors* $u_1, \ldots, u_n \in \mathbb{F}^n$ *with eigenvalues* $\lambda_1, \ldots, \lambda_n$ *if and only if* $A = UDU^{-1}$ *and* $U\overline{U}^t = 1$ *where,*

$$U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} \quad \textit{and} \qquad D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $u_1, \ldots, u_n$ *are the columns of* U *and* D *is the diagonal matrix with diagonal entries* $\lambda_1, \ldots, \lambda_n$.

5 $\mathbb{F}$-modules: Definitions and results

5.1 Vector spaces and linear transformations

Let $\mathbb{F}$ be a field. A $\mathbb{F}$-*vector space*, or $\mathbb{F}$-*module*, is a set V with functions

$$\begin{array}{ccc} V\times V & \to & V \\ (v_1,v_2) & \mapsto & v_1+v_2 \end{array} \qquad \text{and} \qquad \begin{array}{ccc} \mathbb{F}\times V & \to & V \\ (c,v) & \mapsto & cv \end{array}$$

(*addition* and *scalar multiplication*) such that

(a) If $v_1,v_2,v_3\in V$ then $(v_1+v_2)+v_3=v_1+(v_2+v_3)$,
(b) There exists $0\in V$ such that if $v\in V$ then $0+v=v$ and $v+0=v$,
(c) If $v\in V$ then there exists $-v\in V$ such that $v+(-v)=0$ and $(-v)+v=0$,
(d) If $v_1,v_2\in V$ then $v_1+v_2=v_2+v_1$,
(e) If $c\in\mathbb{F}$ and $v_1,v_2\in V$ then $c(v_1+v_2)=cv_1+cv_2$,
(f) If $c_1,c_2\in\mathbb{F}$ and $v\in V$ then $(c_1+c_2)v=c_1v+c_2v$,
(g) If $c_1,c_2\in\mathbb{F}$ and $v\in V$ then $c_1(c_2v)=(c_1c_2)v$,
(h) If $v\in V$ then $1v=v$.

Linear transformations. Linear transformations are for comparing vector spaces.

Let $\mathbb{F}$ be a field and let V and W be $\mathbb{F}$-vector spaces. An $\mathbb{F}$-*linear transformation from V to W* is a function $f\colon V\to W$ such that

(a) If $v_1,v_2\in V$ then $f(v_1+v_2)=f(v_1)+f(v_2)$,
(b) If $c\in\mathbb{F}$ and $v\in V$ then $f(cv)=cf(v)$.

Subspaces. One vector space can be a subspace of another.

Let V be an $\mathbb{F}$-vector space. A *subspace of* V is a subset $W\subseteq V$ such that

(a) If $w_1,w_2\in W$ then $w_1+w_2\in W$,
(b) $0\in W$,
(c) If $w\in W$ then $-w\in W$,
(d) If $w\in W$ and $c\in\mathbb{F}$ then $cw\in W$.

The zero subspace. The tiniest vector space is the zero space.

The *zero space*, (0), is the set containing only 0 with the operations $0+0=0$ and $c\cdot 0$, for $c\in\mathbb{F}$.

5.2 Kernels and images

The *kernel*, or *null space*, of an $\mathbb{F}$-linear transformation $f\colon V\to W$ is the set

$$\ker(f)=\{v\in V\mid f(v)=0\}.$$

The *image* of an $\mathbb{F}$-linear transformation $f\colon V\to W$ is the set

$$\mathrm{im}(f)=\{f(v)\mid v\in V\}.$$

Proposition 5.1. *Let $f\colon V\to W$ be an $\mathbb{F}$-linear transformation. Then* *(a)* $\ker f$ *is a subspace of* V, *and*

(b) $\operatorname{im} f$ *is a subspace of* W.

Let S and T be sets and let $f: S \to T$ be a function.
The function $f: S \to T$ is *injective* if f satisfies:

$$\text{if } s_1, s_2 \in S \text{ and } f(s_1) = f(s_2) \quad \text{then } s_1 = s_2.$$

The function $f: S \to T$ is *surjective* if f satisfies:

$$\text{if } t \in T \text{ then there exists } s \in S \text{ such that } f(s) = t.$$

Proposition 5.2. *Let* $f: V \to W$ *be a linear transformation. Then*

(a) $\ker f = \{0\}$ *if and only if* f *is injective, and*

(b) $\operatorname{im} f = W$ *if and only if* f *is surjective.*

5.3 Bases and dimension

Let $\mathbb{F}$ be a field and let V be a vector space over $\mathbb{F}$. Let $\{v_1, v_2, \ldots, v_k\}$ be a subset of V.

- The *span* of the set $\{v_1, \ldots, v_k\}$ is

$$\operatorname{span}\{v_1, \ldots, v_k\} = \{c_1v_1 + c_2v_2 + \cdots + c_kv_k \mid c_1, c_2, \ldots, c_k \in \mathbb{F}\}.$$

- A *linear combination* of $v_1, v_2, \ldots, v_k$ is an element of $\operatorname{span}\{v_1, \ldots, v_k\}$.

- The set $\{v_1, \ldots, v_k\}$ is *linearly independent* if it satisfies:

$$\text{if } c_1, \ldots, c_k \in \mathbb{F} \text{ and } c_1v_1 + \cdots + c_kv_k = 0 \quad \text{then} \quad c_1 = 0,\ c_2 = 0,\ \ldots,\ c_k = 0.$$

- A *basis* of V is a subset $B \subseteq V$ such that

  (a) $\operatorname{span}(B) = V$,

  (b) B is linearly independent.

- The *dimension* of V is the cardinality (number of elements) of a basis of V.

Proposition 5.3. *Let* V *be a vector space and let* B *be a subset of* V. *The following are equivalent:*

(a) B *is a basis of* V;

(b) B *is a minimal element of* $\{S \subseteq V \mid \operatorname{span}(S) = V\}$, *ordered by inclusion;*

(c) B *is a maximal element of* $\{L \subseteq V \mid L \text{ is linearly independent}\}$, *ordered by inclusion.*

Theorem 5.4. *Let* V *be a vector space over a field* $\mathbb{F}$. *Then*

(a) V *has a basis, and*

(b) Any two bases of V *have the same number of elements.*

5.4 Addition, scalar multiplication and composition of linear transformations

The *sum* of two $\mathbb{F}$-linear transformations $f_1\colon V \to W$ and $f_2\colon V \to W$ is the $\mathbb{F}$-linear transformation $(f_1+f_2)\colon V \to W$.

$$(f_1+f_2)(v) = f_1(v) + f_2(v), \qquad \text{for } v \in V.$$

Let $f\colon V \to W$ be an $\mathbb{F}$-linear transformation and let $c \in \mathbb{F}$. The *scalar multiplication* of f by c is the $\mathbb{F}$-linear transformation $(cf)\colon V \to W$ given by

$$(cf)(v) = c \cdot f(v), \qquad \text{for } v \in V.$$

The *composition* of an $\mathbb{F}$-linear transformation $f_2\colon V \to W$ and an $\mathbb{F}$-linear transformation $f_1\colon W \to Z$ is the $\mathbb{F}$-linear transformation $(f_1 \circ f_2)\colon V \to Z$ given by

$$(f_1 \circ f_2)(v) = f_1(f_2(v)), \qquad \text{for } v \in V.$$

5.5 Matrices of linear transformations and change of basis matrices

Let V and W be $\mathbb{F}$-vector spaces. Let B be a basis of V and let C be a basis of W. A *$C \times B$ matrix with entries in* $\mathbb{F}$ is a function

$$X\colon C \times B \to \mathbb{F}, \qquad \text{where} \quad C \times B = \{(c,b) \mid c \in C \text{ and } b \in B\}.$$

The value $X(c,b)$ of the function X at the pair (c,b) is the *(c,b)-entry of the matrix X*. Denote the set of $C \times B$ matrices with entries in $\mathbb{F}$ by

$$M_{C\times B}(\mathbb{F}) = \{X\colon C \times B \to \mathbb{F}\}.$$

Let $f\colon V \to W$ be an $\mathbb{F}$-linear transformation. The *matrix of $f\colon V \to W$ with respect to the bases B and C* is the matrix

$$f_{CB} \in M_{C\times B}(\mathbb{F}) \qquad \text{given by} \qquad f(b) = \sum_{c\in C} f_{CB}(c,b)c \quad \text{for } b \in B.$$

Proposition 5.5. *Let V and W and Z be $\mathbb{F}$-vector spaces with bases B, C and D, respectively. Let*

$$f\colon V \to W, \quad g\colon V \to W, \quad h\colon W \to Z \quad \text{be } \mathbb{F}\text{-linear transformations}$$

and let $c \in \mathbb{F}$. Then

$$(cf)_{CB} = c \cdot f_{CB}, \qquad f_{CB} + g_{CB} = (f+g)_{CB} \qquad \text{and} \qquad (h \circ g)_{DB} = h_{DC}g_{CB}.$$

Let V be an $\mathbb{F}$-vector space and let B and C be bases of V. The *change of basis matrix from B to C* is the matrix $P_{CB} \in M_{C\times B}(\mathbb{F})$ given by

$$b = \sum_{c\in C} P_{CB}(c,b)c, \qquad \text{for } b \in B. \tag{5.1}$$

Proposition 5.6. *Let $g\colon V \to W$ and $f\colon V \to V$ be an $\mathbb{F}$-linear transformations. Let*

$$B_1 \text{ and } B_2 \text{ be bases of } V, \quad \text{and let} \quad C_1 \text{ and } C_2 \text{ be bases of } W,$$

and let $P_{B_1B_2}$ and $P_{C_2C_1}$ be the change of basis matrices defined as in (5.1). Then

$$g_{C_2B_2} = P_{C_2C_1} g_{C_1B_1} P_{B_1B_2} \qquad \text{and} \qquad f_{B_2B_2} = P^{-1}_{B_1B_2} f_{B_1B_1} P_{B_1B_2}.$$

6 Inner product spaces

6.1 Inner products

Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Let

$$\begin{array}{lll} \overline{}\colon \mathbb{C} \to \mathbb{C} & \text{be given by} & \overline{a+bi} = a - bi, \\ \overline{}\colon \mathbb{R} \to \mathbb{R} & \text{be given by} & \overline{a} = a. \end{array}$$

Let V be an $\mathbb{F}$-vector space. An *inner product on V* is a function

$$\begin{array}{cccc} \langle , \rangle\colon & V \times V & \to & \mathbb{F} \\ & (u,v) & \mapsto & \langle u, v\rangle \end{array} \qquad \text{such that}$$

(1) If $u, v \in V$ then $\langle u, v\rangle = \overline{\langle v, u\rangle}$,
(2) If $u, v \in V$ and $\alpha \in \mathbb{F}$ then $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$,
(3) if $u, v, w \in V$ then $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$,
(4) (Positive semi-definite) If $u \in V$ then $\langle u, u\rangle \in \mathbb{R}_{\geq 0}$.
(5) (Positive definite) If $u \in V$ and $\langle u, u\rangle = 0$ then $u = 0$.

6.2 Length, distance, angles.

Let V be an $\mathbb{F}$-vector space with an inner product.
Length is the function $\| \ \|\colon V \to \mathbb{R}_{\geq 0}$ given by

$$\|u\|^2 = \langle u, u\rangle.$$

Distance is the function $d\colon V \times V \to \mathbb{R}_{\geq 0}$ given by

$$d(u, v) = \|v - u\|.$$

Angle is the function $\theta\colon V \times V \to \mathbb{R}_{[0,\pi]}$ given by

$$\cos(\theta(u, v)) = \frac{\mathrm{Re}(\langle u, v\rangle)}{\|u\| \cdot \|v\|}.$$

Vectors u, v are *orthogonal*, or *perpendicular*, if $\langle u, v\rangle = 0$.
Vectors u, v are *parallel* if $\langle u, v\rangle = 1$ or $\langle u, v\rangle = -1$..

6.3 Standard inner products.

(1) $\begin{array}{cccc} \langle , \rangle\colon & \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ & (u,v) & \mapsto & \langle u|v\rangle \end{array}$ given by

$$\langle u_1, u_2, \ldots, u_n | v_1, v_2, \ldots, v_n\rangle = u_1v_1 + \cdots + u_nv_n,$$

(2) $\begin{array}{cccc} \langle , \rangle\colon & \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ & (u,v) & \mapsto & \langle u|v\rangle \end{array}$ given by

$$\langle u_1, u_2, \ldots, u_n | v_1, v_2, \ldots, v_n\rangle = u_1\overline{v_1} + \cdots + u_n\overline{v_n},$$

(3) $\mathbb{F}[x]_{\leq n} \times \mathbb{F}[x]_{\leq n} \to \mathbb{F}$ given by

$$\begin{aligned} &\langle a_0 + a_1x + \cdots + a_nx^n \mid c_1 + c_1x + \cdots + c_nx^n\rangle \\ &\qquad = \int_0^1 (a_0 + a_1x + \cdots a_nx^n)(\bar{c}_0 + \bar{c}_1x + \cdots + \bar{c}_nx^n)\, dx. \end{aligned}$$

6.4 Gram matrices

Let V be an $\mathbb{F}$-vector space with inner product $\langle,\rangle\colon V\times V\to\mathbb{F}$. Let $B=\{b_1,\ldots,b_n\}$ be a basis of V. The *Gram matrix of* $\langle,\rangle$ *with respect to* B is the matrix

$$A=\begin{pmatrix}\langle b_1,b_1\rangle & \langle b_1,b_2\rangle & \cdots & \langle b_1,b_n\rangle\\ \langle b_2,b_1\rangle & \langle b_2,b_2\rangle & \cdots & \langle b_2,b_n\rangle\\ \vdots & & & \vdots\\ \langle b_n,b_1\rangle & \langle b_n,b_2\rangle & \cdots & \langle b_n,b_n\rangle\end{pmatrix}$$

In other words, the (i,j) entry of the Gram matrix A of $\langle,\rangle$ with respect to the basis B is

$$A_{ij}=\langle b_i,b_j\rangle.$$

6.5 Orthogonal expansions and projections

Let V be an $\mathbb{F}$-vector space with an inner product $\langle,\rangle\colon V\times V\to\mathbb{F}$. A basis $\{b_1,\ldots,b_n\}$ of V is *orthonormal* if $\{b_1,\ldots,b_n\}$ satisfies

$$\text{if } i,j\in\{1,\ldots,n\} \text{ then } \quad \langle b_i,b_j\rangle=\delta_{ij}, \qquad \text{where} \qquad \delta_{ij}=\begin{cases}0, & \text{if } i\neq j,\\ 1 & \text{if } i=j.\end{cases}$$

Proposition 6.1. *Assume $B=(b_1,\ldots,b_n)$ is an ordered orthonormal basis of V and $x\in V$. Then*

$$x=\langle x,b_1\rangle b_1+\cdots+\langle x,b_n\rangle b_n.$$

Let W be a subspace of V. Let $\{b_1,\ldots,b_k\}$ be an orthonormal basis of W. Let $x\in V$. The *orthogonal projection of x onto W* is

$$\mathrm{proj}_W(x)=\langle x,b_1\rangle b_1+\cdots+\langle x,b_k\rangle b_k.$$

6.6 The Gram-Schmidt process of orthogonalization

Let V be an $\mathbb{F}$-vector space with an inner product $\langle,\rangle\colon V\times V\to\mathbb{F}$. An *orthogonal basis* of V is a basis $\{b_1,\ldots,b_n\}$ of V such that

$$\text{if } i,j\in\{1,\ldots,n\} \text{ and } i\neq j \text{ then } \quad \langle b_i,b_j\rangle=0,$$

An *orthonormal basis* of V is a basis $\{b_1,\ldots,b_n\}$ of V such that

$$\text{if } i,j\in\{1,\ldots,n\} \text{ then } \quad \langle b_i,b_j\rangle=\delta_{ij}, \qquad \text{where} \qquad \delta_{ij}=\begin{cases}0, & \text{if } i\neq j,\\ 1 & \text{if } i=j.\end{cases}$$

Theorem 6.2. *(Gram-Schmidt) Let V be an $\mathbb{F}$-vector space with an inner product $\langle,\rangle\colon V\times V\to\mathbb{F}$.*
(a) Let $\{p_1,p_2,\ldots,p_n\}$ be a basis of V. Define $b_1=p_1$ and

$$b_{k+1}=p_{k+1}-\frac{\langle p_{k+1},b_1\rangle}{\langle b_1,b_1\rangle}b_1-\cdots-\frac{\langle p_{k+1},b_k\rangle}{\langle b_k,b_k\rangle}b_k, \qquad \text{for } k\in\mathbb{Z}_{>0}.$$

Then $(b_1,b_2,\ldots,b_n)$ is an orthogonal basis in V.
(b) Let $(b_1,\ldots,b_n)$ be an orthogonal basis of V. Define

$$u_1=\frac{b_1}{\|b_1\|}, \quad \ldots\ , \quad u_n=\frac{b_n}{\|b_n\|}.$$

Then $(u_1,\ldots,u_n)$ is an orthonormal basis of V.

7 Some Proofs

7.1 Some proofs: Matrices and operations

7.1.1 $M_{m\times n}(\mathbb{F})$ is an $\mathbb{F}$-module

Proposition 7.1. *Let $m, n \in \mathbb{Z}_{>0}$ and let $M_{m\times n}(\mathbb{F})$ be the set of $m\times n$ matrices with entries in $\mathbb{F}$.*

(a) If $A, B, C \in M_{m\times n}(\mathbb{F})$ then $A + (B + C) = (A + B) + C$.

(b) If $A, B \in M_{m\times n}(\mathbb{F})$ then $A + B = B + A$.

(c) If $A \in M_{m\times n}(\mathbb{F})$ then $0 + A = A$ and $A + 0 = A$.

(d) If $A \in M_{m\times n}(\mathbb{F})$ then $(-A) + A = 0$ and $A + (-A) = 0$.

(e) If $A \in M_{m\times n}(\mathbb{F})$ and $c_1, c_2 \in \mathbb{F}$ then $c_1 \cdot (c_2 \cdot A) = (c_1c_2) \cdot A$.

(f) If $A \in M_{m\times n}(\mathbb{F})$ and $1 \in \mathbb{F}$ is the identity in $\mathbb{F}$ then $1 \cdot A = A$.

Proof.

(a) To show: $(A + B) + C = A + (B + C)$.
To show: If $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$ then $((A + B) + C)(i, j) = (A + (B + C))(i, j)$.
Assume $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.
To show: $((A + B) + C)(i, j) = (A + (B + C))(i, j)$.

$$\begin{aligned} ((A+B)+C)(i,j) &= (A+B)(i,j) + C(i,j) = (A(i,j) + B(i,j)) + C(i,j) \\ &= A(i,j) + (B(i,j) + C(i,j)), \qquad \text{since addition is associative in } \mathbb{F}, \\ &= A(i,j) + (B+C)(i,j) = (A + (B+C))(i,j). \end{aligned}$$

(b) To show: $A + B = B + A$.
To show: If $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$ then $(A + B)(i, j) = (B + A)(i, j)$.
Assume $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.
To show: $(A + B)(i, j) = (B + A)(i, j)$.

$$\begin{aligned} (A+B)(i,j) &= A(i,j) + B(i,j) \\ &= B(i,j) + A(i,j), \qquad \text{since } \mathbb{F} \text{ has commutative addition}, \\ &= (B+A)(i,j). \end{aligned}$$

(c) To show: (ca) $0 + A = A$.
(cb) $A + 0 = A$.

(ca) To show: If $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$ then $(0 + A)(i, j) = A(i, j)$.
Assume $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.
To show: $(0 + A)(i, j) = A(i, j)$.

$$(0 + A)(i, j) = 0(i, j) + A(i, j) = 0 + A(i, j) = A(i, j).$$

(cb) To show: If $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$ then $(A + 0)(i, j) = A(i, j)$.
Assume $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.

To show: $(A+0)(i,j) = A(i,j)$.

$$(A+0)(i,j) = A(i,j) + 0(i,j) = A(i,j) + 0 = A(i,j).$$

(d) To show: (da) $A + (-A) = 0$.
(db) $(-A) + A = 0$.

(da) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(A + (-A))(i,j) = 0(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(A + (-A))(i,j) = 0(i,j)$.

$$\begin{aligned}(A + (-A))(i,j) &= A(i,j) + (-A)(i,j) = A(i,j) + (-A(i,j)) \\ &= 0 = 0(i,j).\end{aligned}$$

(db) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $((-A) + A)(i,j) = 0(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $((-A) + A)(i,j) = 0(i,j)$.

$$\begin{aligned}((-A + A)(i,j) &= (-A)(i,j) + A(i,j) = (-A(i,j)) + A(i,j) \\ &= 0 = 0(i,j).\end{aligned}$$

(e) Assume $A \in M_{m\times n}(\mathbb{F})$ and $c_1, c_2 \in \mathbb{F}$.
To show $c_1 \cdot (c_2 \cdot A) = (c_1 c_2) \cdot A$.
To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(c_1 \cdot (c_2 \cdot A))(i,j) = ((c_1c_2) \cdot A)(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(c_1 \cdot (c_2 \cdot A))(i,j) = ((c_1c_2) \cdot A)(i,j)$.

$$\begin{aligned}(c_1 \cdot (c_2 \cdot A))(i,j) &= c_1 \cdot ((c_2 \cdot A)(i,j)) = c_1 \cdot (c_2 \cdot A(i,j)) \\ &= (c_1 \cdot c_2) \cdot A(i,j)) = ((c_1 \cdot c_2) \cdot A)(i,j)\end{aligned}$$

(f) Assume $A \in M_{m\times n}(\mathbb{F})$ and $1 \in \mathbb{F}$.
To show: (fa) $1 \cdot A = A$,
(fb) $A \cdot 1 = A$.

(fa) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(1 \cdot A)(i,j) = A(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(1 \cdot A)(i,j) = A(i,j)$.

$$(1 \cdot A)(i,j) = 1 \cdots A(I,j) = A(i,j), \quad \text{since 1 is the identity in } \mathbb{F}.$$

(fb) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(A \cdot 1)(i,j) = A(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(A \cdot 1)(i,j) = A(i,j)$.

$$(A \cdot 1)(i,j) = A(i,j) \cdot 1 = A(i,j), \quad \text{since 1 is the identity in } \mathbb{F}.$$

$\square$

7.1.2 $M_n(\mathbb{F})$ is an $\mathbb{F}$-algebra

Proposition 7.2. *Let $n \in \mathbb{Z}_{>0}$ and let $M_n(\mathbb{F})$ be the set of $n \times n$ matrices in $\mathbb{F}$.*

(a) If $A, B, C \in M_n(\mathbb{F})$ then $A + (B + C) = (A + B) + C$.

(b) If $A, B \in M_n(\mathbb{F})$ then $A + B = B + A$.

(c) If $A \in M_n(\mathbb{F})$ then $0 + A = A$ and $A + 0 = A$.

(d) If $A \in M_n(\mathbb{F})$ then $(-A) + A = 0$ and $A + (-A) = 0$.

(e) If $A, B, C \in M_n(\mathbb{F})$ then $A(BC) = (AB)C$.

(f) If $A, B, C \in M_n(\mathbb{F})$ then $(A + B)C = AC + BC$ and $C(A + B) = CA + CB$.

(g) If $A \in M_n(\mathbb{F})$ then $1A = A$ and $A1 = A$.

Proof. Parts (a), (b) (c) and (d) are the cases $m = n$ of Proposition 1.1.

(e) To show: $(AB)C = A(BC)$.
To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $((AB)C)(i, j) = (A(BC))(i, j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $((AB)C)(i, j) = (A(BC))(i, j)$.

$$\begin{aligned}
((AB)C)(i,j) &= \sum_{k=1}^{n} (AB)(i,k)C(k,j) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n} (A(i,l)B(l,k))C(k,j) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n} A(i,l)(B(l,k)C(k,j)) \\
&= \sum_{l=1}^{n} A(i,l)(BC)(l,j) \\
&= (A(BC))(i,j).
\end{aligned}$$

(f) To show: (fa) $(A + B)C = AC + BC$.
(fb) $C(A + B) = CA + CB$.

(fa) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $((A + B)C)(i, j) = (AC + BC)(i, j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $((A + B)C)(i, j) = (AC + BC)(i, j)$.

$$\begin{aligned}
((A+B)C)(i,j) &= \sum_{k=1}^{n} (A+B)(i,k)C(k,j) = \sum_{k=1}^{n} (A(i,k)+B(i,k))C(k,j) \\
&= \sum_{k=1}^{n} A(i,k)C(k,j) + B(i,k)C(k,j), \\
&= \sum_{k=1}^{n} A(i,k)C(k,j) + \sum_{k=1}^{n} B(i,k)C(k,j), \\
&= (AC)(i,j) + (BC)(i,j) = ((AC)+(BC))(i,j) \\
&= (AC+BC)(i,j).
\end{aligned}$$

where the third equality follows from the distributive property in $\mathbb{F}$.

(fb) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(C(A+B))(i,j) = (CA+CB)(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(C(A+B))(i,j) = (CA+CB)(i,j)$.

$$\begin{aligned}
(C(A+B))(i,j) &= \sum_{k=1}^{n} C(i,k)(A+B)(k,j) = \sum_{k=1}^{n} C(i,k)(A(k,j)+B(k,j)) \\
&= \sum_{k=1}^{n} C(i,k)A(k,j) + C(i,k)B(k,j), \\
&= \sum_{k=1}^{n} C(i,k)A(k,j) + \sum_{k=1}^{n} C(i,k)A(k,j), \\
&= (CA)(i,j) + (CB)(i,j) = ((CA)+(CB))(i,j) \\
&= (CA+CB)(i,j).
\end{aligned}$$

where the third equality follows from the distributive property in $\mathbb{F}$.

(g) To show: (ga) $1 \cdot A = A$,
(gb) $A \cdot 1 = A$.

(ga) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(1 \cdot A)(i,j) = A(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(1 \cdot A)(i,j) = A(i,j)$.

$$\begin{aligned}
(1 \cdot A)(i,j) &= \sum_{k=1}^{n} 1(i,k)A(k,j) = 1(i,i)A(i,j) + \sum_{\substack{k \in \{1,\ldots,n\} \\ k \neq i}} 1(i,k)A(k,j) \\
&= 1 \cdot A(i,j) + 0 = A(i,j).
\end{aligned}$$

(gb) To show: If $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ then $(A \cdot 1)(i,j) = A(i,j)$.
Assume $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
To show: $(A \cdot 1)(i,j) = A(i,j)$.

$$\begin{aligned}
(A \cdot 1)(i,j) &= \sum_{k=1}^{n} A(i,k)1(k,j) = A(i,j)1(j,j) + \sum_{\substack{k \in \{1,\ldots,n\} \\ k \neq j}} A(i,k)1(k,j) \\
&= 0 + A(i,j) \cdot 1 = A(i,j).
\end{aligned}$$

$\square$

7.1.3 Properties of transpose

Proposition 7.3. *Let $m, n \in \mathbb{Z}_{>0}$, let $M_{m\times n}(\mathbb{F})$ be the set of $m \times n$ matrices with entries in $\mathbb{F}$, and let $M_n(\mathbb{F})$ be the set of $n \times n$ matrices in $\mathbb{F}$.*

(a) If $A, B \in M_{m\times n}(\mathbb{F})$ then $(A+B)^T = A^T + B^T$,

(b) If $A \in M_{m\times n}(\mathbb{F})$ and $c \in \mathbb{F}$ then $(c \cdot A)^T = c \cdot A^T$,

(c) If $A, B \in M_n(\mathbb{F})$ then $(AB)^T = B^T A^T$.

(d) If $A \in M_n(\mathbb{F})$ then $(A^T)^T = A$.

Proof.

(a) Assume $A \in M_{m\times n}(\mathbb{F})$.
To show $(A+B)^T = A^T + B^T$.
To show: If $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$ then $(A+B)^T(i,j) = (A^T+B^T)(i,j)$.
Assume $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$.
To show: $(A+B)^T(i,j) = (A^T+B^T)(i,j)$.

$$\begin{aligned}(A+B)^T(i,j) &= (A+B)(j,i) = A(j,i)+B(j,i)\\ &= A^T(i,j)+B^T(i,j) = (A+B)^T(i,j).\end{aligned}$$

(b) Assume $A \in M_{m\times n}(\mathbb{F})$ and $c \in \mathbb{F}$.
To show $(c\cdot A)^T = c\cdot A^T$.
To show: If $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$ then $((c\cdot A)^T)(i,j) = (c\cdot A^T)(i,j)$.
Assume $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$.
To show: $((c\cdot A)^T)(i,j) = (c\cdot A^T)(i,j)$.

$$((c\cdot A)^T)(i,j) = (c\cdot A)(j,i) = c\cdot A(j,i) = c\cdot A^T(i,j) = (c\cdot A^T)(i,j)$$

(c) Assume $A, B \in M_n(\mathbb{F})$.
To show: $(AB)^T = B^T A^T$.
To show: If $i,j \in \{1,\ldots,n\}$ then $(AB)^T(i,j) = (B^TA^T)(i,j)$.
Assume $i,j \in \{1,\ldots,n\}$.
To show: $(AB)^T(i,j) = (B^TA^T)(i,j)$.

$$\begin{aligned}(AB)^T(i,j) = (AB)(j,i) &= \sum_{k=1}^n A(j,k)B(k,i)\\ &= \sum_{k=1}^n A^T(k,j)B^T(i,k) = \sum_{k=1}^n B^T(i,k)A^T(k,j) = (B^TA^T)(i,j).\end{aligned}$$

(d) Assume $A \in M_n(\mathbb{F})$.
To show: $(A^T)^T = A$.
To show: If $i,j \in \{1,\ldots,n\}$ then $((A^T)^T)(i,j) = A(i,j)$.
Assume $i,j \in \{1,\ldots,n\}$.
To show: $((A^T)^T)(i,j) = A(i,j)$.

$$((A^T)^T)(i,j) = (A^T)(j,i) = A(i,j)$$

$\square$

7.1.4 Properties of trace

Proposition 7.4 (Properties of trace). *Let $A, B \in M_{n\times n}(\mathbb{Q})$ and let $c \in \mathbb{Q}$. Then*

$$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B), \qquad \mathrm{tr}(cA) = c\,\mathrm{tr}(A) \qquad \textit{and} \qquad \mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Proof.

$$\begin{aligned}\mathrm{tr}(A+B) &= \sum_{i=1}^{n} (A+B)(i,i) = \sum_{i=1}^{n} (A(i,i) + B(i,i)) \\ &= \Big(\sum_{i=1}^{n} A(i,i)\Big) + \Big(\sum_{i=1}^{n} B(i,i)\Big) = \mathrm{tr}(A) + \mathrm{tr}(B)\end{aligned}$$

and

$$\mathrm{tr}(cA) = \sum_{i=1}^{n} (cA)(i,i) = \sum_{i=1}^{n} (c\cdot A(i,i)) c \cdot \Big(\sum_{i=1}^{n} A(i,i)\Big) = c\,\mathrm{tr}(A)$$

and

$$\begin{aligned}\mathrm{tr}(AB) &= \sum_{i=1}^{n} (AB)(i,i) = \sum_{i=1}^{n}\sum_{j=1}^{n} A(i,j)B(j,i) = \sum_{i=1}^{n}\sum_{j=1}^{n} B(j,i)A(i,j) \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} B(j,i)A(i,j) = \sum_{j=1}^{n} (BA)(j,j) = \mathrm{tr}(BA).\end{aligned}$$

$\square$

7.1.5 Determinant and trace are conjugacy invariants

Theorem 7.5 (Determinant and trace are conjugacy invariants). *Let $n \in \mathbb{Z}_{>0}$. Let $A \in M_{n\times n}(\mathbb{Q})$ and let $P \in GL_n(\mathbb{Q})$ so that P is an invertible $n \times n$ matrix. Then*

$$\det(PAP^{-1}) = \det(A) \qquad \textit{and} \qquad \mathrm{tr}(PAP^{-1}) = \mathrm{tr}(A).$$

Proof. Since $\det(P)\det(P^{-1}) = \det(PP^{-1}) = \det(1) = 1$ then $\det(P^{-1}) = \frac{1}{\det(P)}$. So

$$\det(PAP^{-1}) = \det(P)\det(A)\det(P^{-1}) = \det(P)\det(A)\frac{1}{\det(P)} = \det(A).$$

Since $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ then

$$\mathrm{tr}(PAP^{-1}) = \mathrm{tr}(P(AP^{-1})) = \mathrm{tr}((AP^{-1})P) = \mathrm{tr}(A).$$

$\square$

7.1.6 Determinants: the permutation formula

For $A \in M_n(\mathbb{F})$ define

$$D(A) = \sum_{w\in S_n} \det(w) A(1,w(1))A(2,w(2))\cdots A(n,w(n)),$$

Proposition 7.6. *Let $A \in M_n(\mathbb{F})$ and let $i, j \in \{1, \ldots, n\}$ with $i < j$.*

(a) If u is a permutation then $D(uA) = \det(u)D(A)$.

(b) If row i and row j of A are equal then $D(A) = 0$.

(c) $D(AB) = D(A)D(B)$.

(d) $D(s_{ij}) = -1$, $D(x_{ij}(c)) = 1$ and $D(h_i(d)) = c$.

Proof. Let $u\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a bijection. Then

$$\begin{aligned}
D(uA) &= \sum_{w \in S_n} \det(w) A(u(1), w(1)) A(u(2), w(2)) \cdots A(u(n), w(n)) \\
&= \sum_{w \in S_n} \det(wu) A(u(1), wu(1)) A(u(2), wu(2)) \cdots A(u(n), wu(n)) \\
&= \sum_{w \in S_n} \det(wu) A(1, w(1)) A(2, w(2)) \cdots A(n, w(n)) \\
&= \sum_{w \in S_n} \det(w) \det(u) A(1, w(1)) A(2, w(2)) \cdots A(n, w(n)) \\
&= \det(u) \sum_{w \in S_n} \det(w) A(1, w(1)) A(2, w(2)) \cdots A(n, w(n)) = \det(u) D(A).
\end{aligned}$$

If row i and row j of A are equal then $A = s_{ij}A$. Using part (a),

$$D(A) = D(s_{ij}A) = \det(s_{ij})D(A) = -D(A), \qquad \text{giving} \qquad D(A) = 0. \qquad \text{(Deqrows)}$$

$$\begin{aligned}
D(AB) &= \sum_{w \in S_n} \det(w) (AB)(1, w(1)) \cdots (AB)(n, w(n)) \\
&= \sum_{w \in S_n} \det(w) \sum_{k_1, \ldots, k_n = 1}^{n} A(1, k_1) B(k_1, w(1)) \cdots A(n, k_n) B(k_n, w(n)) \\
&= \sum_{k_1, \ldots, k_n = 1}^{n} A(1, k_1) \cdots A(n, k_n) \sum_{w \in S_n} \det(w) B(k_1, w(1)) \cdots B(k_n, w(n)).
\end{aligned}$$

By part (b), the second sum is zero if any two entries in $k = (k_1, \ldots, k_n)$ are the same. If all the entries in $k = (k_1, \ldots, k_n)$ are distinct and, by part (a), the second sum is $\det(k)D(B)$. Hence,

$$\begin{aligned}
D(AB) &= \sum_{k = (k_1, \ldots, k_n) \in S_n} \det(k) A(1, k_1) \cdots A(n, k_n) \det(k) D(B) \\
&= \sum_{k \in S_n} \det(k) A(1, k(1)) \cdots A(n, k(n)) D(B) = D(A)D(B).
\end{aligned}$$

For each of the expressions (d), the permutation expansion has exactly one nonzero term and this term equals the right hand side. $\square$

Theorem 7.7. *Let $A \in M_n(\mathbb{F})$. Then*

$$\det(A) = \sum_{w \in S_n} \det(w) A(1, w(1)) A(2, w(2) \cdots A(n, w(n)).$$

Proof. By Proposition 7.6, $D(A)$ satisfies the defining properties of $\det(A)$. So $\det(A) = D(A)$. $\square$

7.2 Some proofs: Kernels and images

7.2.1 The inverse of PQ

Proposition 7.8. *If $P, Q \in GL_n(\mathbb{F})$ then*

$$(PQ)^{-1} = Q^{-1}P^{-1}.$$

Proof. Assume $P, Q \in GL_n(\mathbb{F})$.
To show: $(PQ)^{-1} = Q^{-1}P^{-1}$.
To show: (a) $(PQ)(Q^{-1}P^{-1}) = 1$.
(b) $(Q^{-1}P^{-1})(PQ) = 1$.
(a) Using associativity of matrix multiplication,

$$(PQ)(Q^{-1}P^{-1}) = P(QQ^{-1})P^{-1} = P \cdot 1 \cdot P^{-1} = PP^{-1} = 1$$

(b) Using associativity of matrix multiplication,

$$(Q^{-1}P^{-1})(PQ) = Q^{-1}(P^{-1}P)Q = Q^{-1} \cdot 1 \cdot Q = Q^{-1}Q = 1.$$

So $(PQ)^{-1} = Q^{-1}P^{-1}$. □

7.2.2 $\ker(A)$ and $\mathrm{im}(A)$ are subpaces

Proposition 7.9. *Let $A \in M_{t\times s}(\mathbb{Q})$. Then $\ker(A)$ is a subspace of $\mathbb{Q}^s$.*

Proof. (a) Since $A0 = 0$ then $0 \in \ker(A)$.
(b) Assume $w_1, w_2 \in \ker(A)$. Then $Aw_1 = 0$ and $Aw_2 = 0$. So

$$A(w_1 + w_2) = Aw_1 + Aw_2 = 0 + 0 = 0. \quad \text{So } w_1 + w_2 \in \ker(A).$$

(c) Assume $w \in \ker(A)$ and $c \in \mathbb{Q}$. Then $Aw = 0$ and

$$A(cw) = cAw = c0 = 0. \quad \text{So } cw \in \ker(A).$$

So $\ker(A)$ is a subspace of $\mathbb{Q}^s$. □

Proposition 7.10. *Let $A \in M_{t\times s}(\mathbb{Q})$. Then im$(A)$ is a subspace of $\mathbb{Q}^t$.*

Proof. (a) Since $0 = A0$ then $0 \in \mathrm{im}(A)$.
(b)Assume $y_1, y_2 \in \mathrm{im}(A)$. Then there exist $x_1, x_2 \in \mathbb{Q}^s$ such that $y_1 = Ax_1$ and $y_2 = Ax_2$. Then

$$y_1 + y_2 = Ax_1 + Ax_2 = A(x_1 + x_2). \quad \text{So } y_1 + y_2 \in \mathrm{im}(A).$$

(c) Assume $y \in \mathrm{im}(A)$ and $c \in \mathbb{Q}$. Then there exists $x \in \mathbb{Q}^s$ such that $y = Ax$. Then

$$cy = cAx = A(cx). \quad \text{So } cy \in \mathrm{im}(A).$$

So $\mathrm{im}(A)$ is a subspace of $\mathbb{Q}^t$. □

7.2.3 How kernel and image change

Proposition 7.11. ***(How kernel and image change)*** *Let $A \in M_{t\times s}(\mathbb{Q})$ and $P \in GL_t(\mathbb{Q})$ and $Q \in GL_s(\mathbb{Q})$. Then*

$$\begin{aligned} \ker(PA) &= \ker(A), & \ker(AQ) &= Q^{-1}\cdot\ker(A), \\ im(PA) &= P\cdot im(A), & im(AQ) &= im(A). \end{aligned}$$

Proof.

$$\begin{aligned} \ker(PA) &= \{x \in \mathbb{Q}^s \mid PAx = 0\} = \{x \in \mathbb{Q}^s \mid P^{-1}PAx = P^{-1}0\} \\ &= \{x \in \mathbb{Q}^s \mid Ax = 0\} = \ker(A), \\ \ker(AQ) &= \{x \in \mathbb{Q}^s \mid AQx = 0\} = \{Q^{-1}Qx \in \mathbb{Q}^s \mid AQx = 0\} \\ &= Q^{-1}\cdot\{Qx \in \mathbb{Q}^s \mid AQx = 0\} \\ &= Q^{-1}\{y \in \mathbb{Q}^s \mid Ay = 0\} = Q^{-1}\cdot\ker(A), \\ \mathrm{im}(PA) &= \{PAx \mid x \in \mathbb{Q}^s\} = P\{Ax \mid x \in \mathbb{Q}^s\} = P\cdot\mathrm{im}(A), \\ \mathrm{im}(AQ) &= \{AQx \mid x \in \mathbb{Q}^s\} = \{Ay \mid Q^{-1}y \in \mathbb{Q}^s\} \\ &= \{Ay \mid y \in \mathbb{Q}^s\} = \mathrm{im}(A). \qquad \square \end{aligned}$$

7.2.4 Finding $\ker(A)$ and $\mathrm{im}(A)$

Theorem 7.12. ***(Computing kernels and images)*** *Let $P \in GL_t(\mathbb{Q})$, $Q \in GL_s(\mathbb{Q})$. Let $r \in \{1,\ldots,\min(s,t)\}$ and let*

$$1_r = E_{11} + E_{22} + \cdots + E_{rr} \quad in \quad M_{t\times s}(\mathbb{Q}).$$

Let

$$A = P1_rQ.$$

Then

$$\ker(A)\ \textit{has basis}\ \{\textit{last}\ s{-}r\ \textit{columns of}\ Q^{-1}\},$$

$$im(A)\ \textit{has basis}\ \{\textit{first}\ r\ \textit{columns of}\ P\}.$$

Proof. By the How Kernel and Image Change Proposition

$$\begin{aligned} \ker(A) &= \ker(P1_rQ) = \ker(1_rQ) = Q^{-1}\ker(1_r) \quad \text{and} \\ \mathrm{im}(A) &= \mathrm{im}(P1_rQ) = \mathrm{im}(P1_r) = P\mathrm{im}(1_r) \end{aligned}$$

Since $\{Q^{-1}e_{r+1},\ldots,Q^{-1}e_s\}$ is a basis of $Q^{-1}\ker(1_r)$ then

$$\ker(A)\ \text{has basis}\ \{\text{last}\ s{-}r\ \text{columns of}\ Q^{-1}\},$$

Since $\{Pe_1,\ldots,Pe_r\}$ is a basis of $P\mathrm{im}(1_r)$ then

$$\mathrm{im}(A)\ \text{has basis}\ \{\text{first}\ r\ \text{columns of}\ P\}. \qquad \square$$

7.2.5 Relating kernels and determinants

Theorem 7.13. *Let $A \in M_n(\mathbb{F})$ and let $\lambda \in \mathbb{F}$. Then*

$$\ker(A-\lambda) \neq 0 \qquad \text{if and only if} \qquad \det(A-\lambda) = 0.$$

Proof. Use normal form to write $A - \lambda = P \cdot 1_r \cdot Q$ with $P, Q \in GL_n(\mathbb{F})$. Then use Proposition **??**.
$\Rightarrow$: Assume $\ker(A-\lambda) \neq 0$.
Since $\ker(A-\lambda) \neq 0$ then $r < n$.
Since $r < n$ then $\det(1_r) = 0$ and

$$\det(A-\lambda) = \det(P \cdot 1_r \cdot Q) = \det(P)\det(1_r)\det(Q) = 0.$$

$\Leftarrow$: Assume that $\det(A-\lambda) = 0$.
Since P and Q are invertible matrices then $\det(P)$ and $\det(Q)$ are invertible elements of $\mathbb{F}$.
Since

$$\det(1_r) = \det(P^{-1}(A-\lambda)Q^{-1}) = \det(P)^{-1}\det(A-\lambda)\det(Q)^{-1} = \det(P)^{-1} \cdot 0 \cdot \det(Q)^{-1} = 0$$

then $r < n$.
So $\ker(1_r) \neq 0$.
So

$$\ker(A-\lambda) = \ker(P \cdot 1_r \cdot Q) = Q^{-1}\ker(1_r) \neq 0.$$

□

7.2.6 Solutions of a system when A is invertible

Proposition 7.14. *Let $A \in M_{t\times s}(\mathbb{F})$ and $b \in \mathbb{F}^t$. If $t = s$ and $A \in GL_s(\mathbb{F})$ then*

$$\mathrm{Sol}(Ax = b) = \{A^{-1}b\}.$$

Proof. Assume A is invertible and $Ax = b$. Then $x = 1x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}b$. □

7.2.7 All solutions from a single solution and $\ker(A)$

Proposition 7.15. *Let $A \in M_{t\times s}(\mathbb{F})$ and $b \in \mathbb{F}^t$ and assume $\mathrm{Sol}(Ax = b) \neq \emptyset$. Let $p \in \mathrm{Sol}(Ax = b)$. Then*

$$\mathrm{Sol}(Ax = b) = p + \ker(A).$$

Proof. Assume $\mathrm{Sol}(Ax = b) \neq \emptyset$ and let $p \in \mathrm{Sol}(Ax = b)$.
To show: (a) $\mathrm{Sol}(Ax = b) \subseteq p + \ker(A)$.
(b) $p + \ker(A) \subseteq \mathrm{Sol}(ax = b)$.

(a) Let $q \in Sol(Ax = b)$.
Then $q - p \in \ker(A)$.
So $q \in p + \ker(A)$.
So $Sol(Ax = b) \subseteq p + \ker(A)$.

[(b) Let $k \in \ker(A)$.
Then $A(p + k) = Ap + Ak = b + 0 = b$.
So $p + k \in p + \ker(A)$.
So $p + \ker(A) \subseteq \mathrm{Sol}(Ax = b)$.

So $p + \ker(A) = \mathrm{Sol}(Ax = b)$ □

7.2.8 All solutions of a system $Ax = b$

Proposition 7.16. *Let $A \in M_{t\times s}(\mathbb{F})$ and $b \in \mathbb{F}^t$. Assume $r \in \{1, \ldots, \min(s,t)\}$ and $P \in GL_t(\mathbb{F})$ and $Q \in GL_s(\mathbb{F})$ are such that*

$$A = P1_rQ.$$

(a) If there exists $j \in \{r+1, \ldots, t\}$ such that $(P^{-1}b)_j \neq 0$ then $\mathrm{Sol}(Ax = b) = \emptyset$.

(b) If $(P^{-1}b)_{r+1} = 0, (P^{-1}b)_{r+2} = 0, \ldots, (P^{-1}b)_t = 0$ then $\mathrm{Sol}(Ax = b) \neq \emptyset$ and

$$\mathrm{Sol}(Ax = b) = Q^{-1}((P^{-1}b)_1, \ldots, (P^{-1}b)_r, 0, \ldots, 0)^t + \mathrm{span}\{q_{r+1}, \ldots, q_s\},$$

where $q_1, \ldots, q_s$ are the columns of Q^{-1}.

Proof.

$$\begin{aligned}
\mathrm{Sol}(Ax = b) &= \{x \in \mathbb{F}^n \mid Ax = b\} = \{x \in \mathbb{F}^n \mid P1_rQx = b\} \\
&= Q^{-1}\{Qx \in \mathbb{F}^n \mid P1_rQx = b\} \\
&= Q^{-1}\{y \in \mathbb{F}^n \mid P1_ry = b\} \\
&= Q^{-1}\{y \in \mathbb{F}^n \mid 1_ry = P^{-1}b\} \\
&= Q^{-1}\mathrm{Sol}(1_ry = P^{-1}b).
\end{aligned}$$

Let $z = P^{-1}b$. There are two cases:
Case 1. If there exists $j \in \{r+1, \ldots, n\}$ such that $z_j \neq 0$ then $\mathrm{Sol}(1_ry = z) = \emptyset$.
Case 2. If $z_{r+1} = \cdots = z_t = 0$ then

$$\mathrm{Sol}(1_ry = z) = (z_1, \ldots, z_r, 0, \ldots, 0)^t + \mathrm{span}\{e_{r+1}, \ldots e_s\}.$$

Thus

$$\begin{aligned}
\mathrm{Sol}(Ax = b) &= Q^{-1}\mathrm{Sol}(1_ry = P^{-1}b) = Q^{-1}\mathrm{Sol}(1_ry = z) \\
&= Q^{-1}(z_1, \ldots, z_r, 0, \ldots, 0)^t + Q^{-1}\mathrm{span}\{e_{r+1}, \ldots e_s\} \\
&= Q^{-1}((P^{-1}b)_1, \ldots, (P^{-1}b)_r, 0, \ldots, 0)^t + \mathrm{span}\{q_{r+1}, \ldots, q_s\}.
\end{aligned}$$

□

7.2.9 Diagonalization and eigenvectors

Theorem 7.17. *Let $n \in \mathbb{Z}_{>0}$ and let $A \in M_n(\mathbb{F})$.*
The matrix A has n linearly independent eigenvectors $p_1, \ldots, p_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if $A = PDP^{-1}$ where,

$$P = \begin{pmatrix} | & & | \\ p_1 & \cdots & p_n \\ | & & | \end{pmatrix} \qquad \text{and} \qquad D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $p_1, \ldots, p_n$ are the columns of P and D is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$.

Proof. $\Rightarrow$: Assume that $A \in M_n(\mathbb{F})$ and A has n linearly independent eigenvectors $p_1, \ldots, p_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.
Since $Ap_1 = \lambda_1 p_1, \ldots, Ap_n = \lambda_n p_n$ then

$$AP = A\begin{pmatrix} | & & | \\ p_1 & \cdots & p_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ p_1 & \cdots & p_n \\ | & & | \end{pmatrix}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = PD.$$

Since $p_1, \ldots, p_n$ are linearly independent then the matrix P is invertible.
Multiplying $AP = PD$ on the right by P^{-1} gives $A = PDP^{-1}$.

$\Leftarrow$: Assume that $A = PDP^{-1}$ and that $p_1, \ldots, p_n$ are the columns of P.
Then $AP = DP$ gives that $Ap_1 = \lambda_1 p_1, \ldots, Ap_n = \lambda_n p_n$.
So $p_1, \ldots, p_n$ are eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$.
Since $p_1, \ldots, p_n$ are the columns of P and P is invertible then $p_1, \ldots, p_n$ are linearly independent. $\square$

7.3 Some proofs: $\mathbb{F}$-modules

7.3.1 Properties of linear transformations

Proposition 7.18. *Let $T\colon V \to W$ be an $\mathbb{F}$-linear transformation. Let 0_V and 0_W be the zeros for V and W respectively. Then*

(a) $T(0_V) = 0_W$, *and*

(b) If $v \in V$ then $T(-v) = -T(v)$.

Proof.

(a) Add $-T(0_V)$ to both sides of the following equation,

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V).$$

(b) Since $T(v) + T(-v) = T\big(v + (-v)\big) = T(0_V) = 0_W$ and

$$T(-v) + T(v) = T\big((-v) + v\big) = T(0_V) = 0_W$$

then $-T(v) = T(-v)$.

□

7.3.2 Kernels and images are subspaces

Proposition 7.19. *Let $T\colon V \to W$ be an $\mathbb{F}$-linear transformation. Then*

(a) $\ker T$ *is a subspace of* V.

(b) $\mathrm{im}\ T$ *is a subspace of* W.

Proof. Let 0_V and 0_W be the zeros in V and W, respectively.

(a) By condition (a) in the definition of linear transformation, T is a group homomorphism.
To show: (aa) If $k_1, k_2 \in \ker T$ then $k_1 + k_2 \in \ker T$.
(ab) $0_V \in \ker T$.
(ac) If $k \in \ker T$ then $-k \in \ker T$.
(ad) If $c \in \mathbb{F}$ and $k \in \ker T$ then $ck \in \ker T$.

(aa) Assume $k_1, k_2 \in \ker T$.
Then $T(k_1) = 0_W$ and $T(k_2) = 0_W$.
By condition (a) in the definition of a linear transformation,

$$T(k_1 + k_2) = T(k_1) + T(k_2) = 0 + 0 = 0.$$

So $k_1 + k_2 \in \ker T$.

(ab) By Proposition 7.18(a), $T(0_V) = 0_W$.
So $0_V \in \ker T$.

(ac) Assume $k \in \ker T$.
By Proposition 7.18(b), $T(-k) = -T(k)$.
So $T(-k) = -T(k) = -0_W = 0_W$, and $-0_W = 0_W$ since $0_W + 0_W = 0_W$.
So $-k \in \ker T$.

(ad) Assume $c \in \mathbb{F}$ and $k \in \ker T$.
Then, by the definition of linear transformation,

$$T(ck) = cT(k) = c\,0_W = 0_W, \qquad \text{and} \quad c\,0_W = 0_W,$$

by adding $-c\,0_W$ to each side of $c\,0_W + c\,0_W = c(0_W + 0_W) = c\,0_W$.
So $T(ck) = 0_W$ and $ck \in \ker T$.

So $\ker T$ is a subspace of V.

(b) By condition (a) in the definition of an $\mathbb{F}$-linear transformation, T is a group homomorphism.
To show: (ba) If $w_1, w_2 \in \mathrm{im}\,T$ then $w_1 + w_2 \in \mathrm{im}\,T$.
(bb) $0_W \in \mathrm{im}\,T$.
(bc) If $w \in \mathrm{im}\,T$ then $-w \in \mathrm{im}\,T$.
(bd) If $c \in \mathbb{F}$ and $w \in \mathrm{im}\,T$ then $ck \in \mathrm{im}\,T$.

(ba) Assume $w_1, w_2 \in \mathrm{im}\,T$.
Then there exist $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$.
By condition (a) in the definition of an $\mathbb{F}$-linear transformation,

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2.$$

So $w_1 + w_2 \in \mathrm{im}\,T$.
(bb) By Proposition 7.18(a), $T(0_V) = 0_W$.
So $0_W \in \mathrm{im}\,T$.
(bc) Assume $w \in \mathrm{im}\,T$.
The there exists $v \in V$ such that $T(v) = w$.
By Proposition 7.18(b), $T(-v) = -T(v) = -w$.
So $-w \in \mathrm{im}\ T$.
(bd) To show: If $c \in \mathbb{F}$ and $a \in \mathrm{im}\,T$ then $ca \in \mathrm{im}\,T$.
Assume $c \in \mathbb{F}$ and $c \in \mathrm{im}\,T$.
Then there exists $v \in V$ such that $a = T(v)$.
By the definition of an $\mathbb{F}$-linear transformation,

$$ca = cT(v) = T(cv).$$

So $ca \in \mathrm{im}\,T$.

So $\mathrm{im}\,T$ is a subspace of W.

□

7.3.3 Injectivity and surjectivity in terms of kernels and images

Proposition 7.20. *Let $T\colon V \to W$ be an $\mathbb{F}$-linear transformation. Let 0_V be the zero in V. Then*

(a) $\ker T = (0_V)$ *if and only if T is injective.*
(b) $\mathrm{im} T = W$ *if and only if T is surjective.*

Proof. Let 0_V and 0_W be the zeros in V and W respectively.

(a) $\Longrightarrow$: Assume $\ker T = (0_V)$.
To show: If $T(v_1) = T(v_2)$ then $v_1 = v_2$.
Assume $T(v_1) = T(v_2)$.
Since T is an $\mathbb{F}$-linear transformation then

$$0_W = T(v_1) - T(v_2) = T(v_1 - v_2).$$

So $v_1 - v_2 \in \ker T$.
Since $\ker T = (0_V)$ then $v_1 - v_2 = 0_V$.
So $v_1 = v_2$.
So T is injective.

$\Longleftarrow$: Assume T is injective
To show: (aa) $(0_V) \subseteq \ker T$.
(ab) $\ker T \subseteq (0_V)$.

(aa) Since $T(0_V) = 0_W$ then $0_V \in \ker T$.
So $(0_V) \subseteq \ker T$.
(ab) Let $k \in \ker T$.
Then $T(k) = 0_W$.
So $T(k) = T(0_V)$.
Thus, since T is injective then $k = 0_V$.
So $\ker T \subseteq (0_V)$.

So $\ker T = (0_V)$.

(b) $\Longrightarrow$: Assume $\operatorname{im} T = W$.
To show: If $w \in W$ then there exists $v \in V$ such that $T(v) = w$.
Assume $w \in W$.
Then $w \in \operatorname{im} T$.
So there exists $v \in V$ such that $T(v) = w$.
So T is surjective.

$\Longleftarrow$: Assume T is surjective.
To show: (ba) $\operatorname{im} T \subseteq W$.
(bb) $W \subseteq \operatorname{im} T$.

(ba) Let $x \in \operatorname{im} T$.
Then there exists $v \in V$ such that $x = T(v)$.
By the definition of T, $T(v) \in W$.
So $x \in W$.
So $\operatorname{im} T \subseteq W$.

(bb) Assume $x \in W$.
Since T is surjective there exists $v \in V$ such that $T(v) = x$.
So $x \in \operatorname{im} T$.
So $W \subseteq \operatorname{im} T$.

So $\operatorname{im} T = W$.

$\square$

7.3.4 Minimal spanning sets and maximal linearly independent sets

Proposition 7.21. *Let V be an $\mathbb{F}$-vector space and let B be a subset of V. The following are equivalent:*

(a) B is a basis of V.
(b) B is a minimal element of $\{S \subseteq V \mid span_{\mathbb{F}}(S) = V\}$.
(c) B is a maximal element of $\{L \subseteq V \mid L \text{ is linearly independent}\}$.

(In (b) and (c) the ordering is by inclusion.)

Proof.
(b) $\Rightarrow$ (a): Let $S \subseteq V$ such that $\mathrm{span}_{\mathbb{F}}(S) = V$.

To show: If S is minimal such that $\mathrm{span}_{\mathbb{F}}(V)$ then S is a basis.
To show: If S is minimal such that $\mathrm{span}_{\mathbb{F}}(V)$ then S is linearly independent.
Proof by contrapositive.
To show: If S is not linearly independent then S is not minimal such that $\mathrm{span}_{\mathbb{F}}(S) = V$.
Assume S is not linearly independent.
To show: There exists $s \in S$ such that $\mathrm{span}_{\mathbb{F}}(S - \{s\}) = V$.
Since S is linearly independent then there exist $k \in \mathbb{Z}_{>0}$ and $s_1, \ldots, s_k \in S$ and $c_1, \ldots, c_k \in \mathbb{F}$ and $i \in \{1, \ldots, k\}$ such that $c_1 s_1 + \cdots + c_k s_k = 0$ and $c_i \neq 0$.
Let $s = s_i$.
Using that $\mathbb{F}$ is a field and $c_i \neq 0$ then

$$\begin{aligned} s = s_i &= c_i^{-1}(c_1 s_1 + \ldots + c_{i-1}s_{i-1} + c_{i+1}s_{i+1} + \cdots + s_k c_k) \\ &= c_i^{-1}c_1 s_1 + \cdots + c_i^{-1}c_{i-1}s_{i-1} + c_i^{-1}c_{i+1}s_{i+1} + \cdots + c_i^{-1}c_k s_k. \end{aligned}$$

So $V = \mathrm{span}_{\mathbb{F}}(S) = \mathrm{span}_{\mathbb{F}}(S - \{s\})$.
So S is not minimal such that $\mathrm{span}_{\mathbb{F}}(S) = V$.

(a) $\Rightarrow$ (b): Proof by contrapositive.

To show: If $B \in \{S \subseteq V \mid \mathrm{span}_{\mathbb{F}}(S) = V\}$ and B is not minimal then B is not a basis of V.
Assume $B \in \{S \subseteq V \mid \mathrm{span}_{\mathbb{F}}(S) = V\}$ is not minimal.
So there exists $b \in B$ such that $\mathrm{span}_{\mathbb{F}}(B - \{b\}) = V$.
To show: B is not linearly independent.
Since $b \in V = \mathrm{span}(B - \{b\}$ then there exist $k \in \mathbb{Z}_{>0}$, $b_1, \ldots, b_k \in B$ and $c_1, \ldots, c_k \in \mathbb{F}$ such that $b = c_1 b_1 + \cdots + c_k b_k$.
So $0 = c_1 b_1 + \cdots + c_k b_k + (-1)b$ and B is not linearly independent.

(a) $\Rightarrow$ (c): Assume B is a basis of V.

Since B is linearly independent then $B \in \{L \subseteq V \mid L \text{ is linearly independent}\}$.
To show: If $v \in V$ and $v \notin B$ then $B \cup \{v\}$ is not linearly independent.
Assume $v \in V$ and $v \notin B$.
Since $\mathrm{span}_{\mathbb{F}}(B) = V$ then there exists $k \in \mathbb{Z}_{>0}$ and $b_1, \ldots, b_k \in B$ and $c_1, \ldots, c_k \in \mathbb{F}$ such that $v = c_1 b_1 + \ldots + c_k b_k$.
So $0 = c_1 b_1 + \cdots + c_k b_k + (-1)v$.

So $B \cup \{v\}$ is not linearly independent.

(c) $\Rightarrow$ (a): Assume S is a maximal element of $\{L \subseteq V \mid L \text{ is linearly independent}\}$.

To show: $\mathrm{span}_{\mathbb{F}}(S) = V$.
To show: $V \subseteq \mathrm{span}_{\mathbb{F}}(S)$.
Let $v \in V$.
To show: $v \in \mathrm{span}_{\mathbb{F}}(S)$.
Case 1: $v \in S$. Then $v \in \mathrm{span}_{\mathbb{F}}(S)$.
Case 2: $v \notin S$.
Then $S \cup \{v\}$ is not linearly independent and S is linearly independent.
So there exist $k \in \mathbb{Z}_{>0}$ and $s_1, \ldots, s_k \in S$ and $c_0, c_1, \ldots, c_k \in \mathbb{F}$ such that

$$c_0 \neq 0 \qquad \text{and} \qquad c_0 v + c_1 s_1 + \cdots + c_k s_k = 0.$$

Since $\mathbb{F}$ is a field and $c_0 \neq 0$ then

$$v = (-c_0^{-1} c_1) s_1 + \cdots + (-c_0^{-1} c_k) s_k.$$

So $v \in \mathrm{span}_{\mathbb{F}}(S)$.
So $V \subseteq \mathrm{span}_{\mathbb{F}}(S)$ and $V = \mathrm{span}_{\mathbb{F}}(S)$.
So S is linearly independent and $\mathrm{span}_{\mathbb{F}}(S) = V$.
So S is a basis of V.

$\square$

7.3.5 Any two bases have the same number of elements

Theorem 7.22. *Let V be an $\mathbb{F}$-vector space. Then*

(a) V has a basis, and
(b) Any two bases of V have the same number of elements.

Proof.

(a) The idea is to use Zorn's lemma on the set $\{L \subseteq V \mid L \text{ is linearly independent}\}$, ordered by inclusion. We will not prove Zorn's lemma, we will assume it. Zorn's lemma is equivalent to the axiom of choice. For a proof see Isaacs book [**?**, §11D].

> **Zorn's Lemma.** *If S is a nonempty poset such that every chain in S has an upper bound then S has a maximal element.*

Let $v \in V$ such that $v \neq 0$.
Then $L = \{v\}$ is linearly independent.
So $\{L \subseteq V \mid L \text{ is linearly independent}\}$ is not empty.
To show: If $\cdots \subseteq S_{k-1} \subseteq S_k \subseteq S_{k+1} \subseteq \cdots$ chain of linearly independent subsets of V then there exists a linearly independent set S that contains all the S_k.
Assume $\cdots \subseteq S_{k-1} \subseteq S_k \subseteq S_{k+1} \subseteq \cdots$ is a chain of linearly independent subsets of V.

Let $L = \bigcup_k S_k$.
To show L is linearly independent.
Assume $\ell \in \mathbb{Z}_{>0}$ and $s_1, \ldots, s_\ell \in L$.
Then there exists k such that $s_1, \ldots, s_\ell \in S_k$.
Since S_k is linearly independent then if $c_1, \ldots, c_\ell \in \mathbb{F}$ and $c_1 s_1 + \cdots + c_\ell s_\ell = 0$ then $c_1 = 0$, $c_2 = 0, \ldots, c_\ell = 0$.
So L is linearly independent.
So, if $\cdots \subseteq S_{k-1} \subseteq S_k \subseteq S_{k+1} \subseteq \cdots$ chain of linearly independent subsets of V then there exists a linearly independent set B that contains all the S_k.
Thus, by Zorn's lemma, $\{L \subseteq V \mid L \text{ is linearly independent}\}$ has a maximal element B.
By Proposition 5.3, B is a basis of V.

(b) Let B and C be bases of V.

Case 1: V has a basis B with $\mathrm{Card}(B) < \infty$.
Let $b \in B$.
Then there exists $c \in C$ such that $c \notin \mathrm{span}_{\mathbb{F}}(B - \{b\})$.
Then $B_1 = (B - \{b\}) \cup \{c\}$ is a basis with the same cardinality as B.
Since B is finite then, by repeating this process, we can, after a finite number of steps, create a basis B' of V such that $B' \subseteq C$ and $\mathrm{Card}(B') = \mathrm{Card}(B)$.
Thus $\mathrm{Card}(B) = \mathrm{Card}(B') \le \mathrm{Card}(C)$.
A similar argument with C in place of B gives that $\mathrm{Card}(B) \ge \mathrm{Card}(C)$.
So $\mathrm{Card}(B) = \mathrm{Card}(C)$.

Case 2: V has an infinite basis B.
Let C be a basis of V.
Define $P_{cb} \in \mathbb{F}$ for $c \in C$ and $b \in B$ by

$$b = \sum_{c \in C} P_{cb} c, \quad \text{and let} \qquad S_b = \{c \in C \mid P_{cb} \neq 0\} \quad \text{for } b \in B.$$

If $b \in B$ then S_b is a finite subset of C and

$$C = \bigcup_{b \in B} S_b, \qquad \text{since } C \text{ is a minimal spanning set.}$$

So $\mathrm{Card}(C) \le \max\{\mathrm{Card}(S_b) \mid b \in B\} \le \aleph_0 \mathrm{Card}(B)$.
A similar argument with B and C switched shows that $\mathrm{Card}(B) \le \aleph_0 \mathrm{Card}(C)$.
So $\mathrm{Card}(C) \le \aleph_0 \mathrm{Card}(B) = \mathrm{Card}(B) \le \aleph_0 \mathrm{Card}(C) = \mathrm{Card}(C)$.
Since $\mathrm{Card}(C) \le \mathrm{Card}(B) \le \mathrm{Card}(C)$ then $\mathrm{Card}(C) = \mathrm{Card}(B)$.

□

7.3.6 Matrix of a linear transformation with respect to a basis

Proposition 7.23. *Let V and W and Z be $\mathbb{F}$-vector spaces with bases B, C and D, respectively. Let*

$$f\colon V \to W, \quad g\colon V \to W, \quad h\colon W \to Z \quad \textit{be linear transformations}$$

and let $c \in \mathbb{F}$. Then

$$(cf)_{CB} = c \cdot f_{CB}, \qquad f_{CB} + g_{CB} = (f+g)_{CB} \qquad \textit{and} \qquad (h \circ g)_{DB} = h_{DC} g_{CB}.$$

Proof. Let $b \in B$ and $c' \in C$. Taking the coefficient of c' on each side of

$$\sum_{c\in C}(\alpha f)_{CB}(c,b)c = (\alpha f)(b) = \alpha\cdot f(b) = \alpha\cdot\Big(\sum_{c\in C} f_{CB}(c,b)c\Big) = \sum_{c\in C}\alpha f_{CB}(c,b)c$$

gives $(\alpha f)_{CB}(c',b) = \alpha\cdot f_{CB}(c',b)$.
So $(\alpha f)_{CB} = \alpha\cdot f_{CB}$.
Let $b \in B$ and $c' \in C$. Taking the coefficient of c' on each side of

$$\begin{aligned}\sum_{c\in C}(f+g)_{CB}(c,b)c = (f+g)(b) = f(b)+g(b) &= \sum_{c\in C}(f_{CB}(c,b)c + \sum_{c\in C} g_{CB}(c,b)c \\ &= \sum_{c\in C}(f_{CB}(c,b)c + g_{CB}(c,b)c = \sum_{c\in C}(f_{CB}(c,b)+g_{CB}(c,b))c\end{aligned}$$

gives $(f_{CB}+g_{CB})(c',b) = f_{CB}(c',b)+g_{CB}(c',b)$.
So $f_{CB}+g_{CB} = (f+g)_{CB}$.
Let $b \in B$ and $d' \in D$. Taking the coefficient of d' on each side of

$$\begin{aligned}\sum_{d\in D}(h\circ g)_{DB}(d,b)d = (h\circ g)(b) = h(g(b)) &= h\Big(\sum_{c\in C} g_{CB}(c,b)c\Big) \\ &= \sum_{c\in C} g_{CB}(c,b)h(c) = \sum_{c\in C}\sum_{d\in D} g_{CB}(c,b)h_{DC}(d,c)d,\end{aligned}$$

gives $(h\circ g)_{DB}(d',b) = \displaystyle\sum_{c\in C}\sum_{d\in D} h_{DC}(d,'c)g_{CB}(c,b) = (h_{DC}g_{CB})(d',b)$.
So $(h\circ g)_{DB} = (h_{DC}g_{CB})$. □

7.3.7 Change of basis matrices

Proposition 7.24. *Let $g\colon V\to W$ and $f\colon V\to V$ be $\mathbb{F}$-linear transformations. Let*

$$B_1 \text{ and } B_2 \text{ be bases of } V, \quad \text{and let} \quad C_1 \text{ and } C_2 \text{ be bases of } W,$$

and let $P_{B_1B_2}$ and $P_{C_2C_1}$ be the change of basis matrices defined as in (5.1). Then

$$g_{C_2B_2} = P_{C_2C_1}g_{C_1B_1}P_{B_1B_2} \qquad \text{and} \qquad f_{B_2B_2} = P_{B_1B_2}^{-1}f_{B_1B_1}P_{B_1B_2}.$$

Proof. Let $\beta,\beta' \in B_2$. Comparing coefficients of β' on each side of

$$\begin{aligned}\beta &= \sum_{b\in B_1} P_{B_1B_2}(b,\beta)b = \sum_{b\in B_1} P_{B_1B_2}(b,\beta)\sum_{\beta'\in B_2} P_{B_2B_1}(\beta',b)\beta' \\ &= \sum_{b\in B_1}\sum_{\beta'\in B_2} P_{B_2B_1}(\beta',b)P_{B_1B_2}(b,\beta)\beta' = \sum_{b\in B_1}\sum_{\beta'\in B_2}(P_{B_2B_1}P_{B_1B_2})(\beta',\beta)\beta'\end{aligned}$$

gives

$$(P_{B_2B_1}P_{B_1B_2})(\beta',\beta) = \delta_{\beta'\beta}.$$

So $P_{B_2B_1} = P_{B_1B_2}^{-1}$.
Let $\beta \in B_1$ and $c \in B_2$. Taking the coefficient of b' on each side of

$$f(c) = \sum_{c'\in B_2} f_{B_2B_2}(c',c)c' = \sum_{b'\in B_1} f_{B_2B_2}(c',c)P_{B_1B_2}(b',c')b'$$

and

$$f(c) = f\Big(\sum_{b\in B_1} P_{B_1B_2}(b,c)b\Big) = \sum_{b\in B_1} P_{B_1B_2}(b,c)f(b) = \sum_{b\in B_1} P_{B_1B_2}(b,c) \sum_{b'\in B_1} f_{B_1B_1}(b',b)b'$$

gives

$$(P_{B_1B_2}f_{B_2B_2})(\beta,b) = (f_{B_1B_1}P_{B_1B_2})(\beta,b).$$

So

$$P_{B_1B_2}f_{B_2B_2} = f_{B_1B_1}P_{B_1B_2} \qquad \text{and thus} \qquad f_{B_2B_2} = P_{B_1B_2}^{-1}f_{B_1B_1}P_{B_1B_2}.$$

Let $\gamma' \in C_2$ and $\beta \in B_2$. Taking the coefficient of γ on each side of

$$\begin{aligned}
\sum_{\gamma\in C_2} g_{C_2B_2}(\gamma,\beta)\gamma = g(\beta) = g(\sum_{b\in B_1} P_{B_1B_2}(b,\beta)b) &= \sum_{b\in B_1} P_{B_1B_2}(b,\beta)g(b) \\
&= \sum_{b\in B_1} P_{B_1B_2}(b,\beta)\sum_{c\in C_1} g_{C_1B_1}(c,b)c \\
&= \sum_{b\in B_1} P_{B_1B_2}(b,\beta)\sum_{c\in C_1} g_{C_1B_1}(c,b)\sum_{\gamma\in C_2} P_{C_2C_1}(\gamma,c)\gamma \\
&= \sum_{b\in B_1, c\in C_1, \gamma\in C_2} P_{C_2C_1}(\gamma,c)g_{C_1B_1}(c,b)P_{B_1B_2}(b,\beta)\gamma \\
&= \sum_{\gamma\in C_2} (P_{C_2C_1}g_{C_1B_1}P_{B_1B_2})(\gamma,\beta)\gamma
\end{aligned}$$

gives $g_{C_2B_2}(\gamma',\beta) = (P_{C_2C_1}g_{C_1B_1}P_{B_1B_2})(\gamma',\beta)$. So $g_{C_2B_2} = P_{C_2C_2}g_{C_1B_1}P_{B_1B_2}$. □

7.3.8 Invertiblity and column linear independence

Proposition 7.25. *Let $P \in M_n(\mathbb{F})$. The matrix P is invertible if and only if the columns of P are linearly independent in $\mathbb{F}^n$.*

Proof.
$\Rightarrow$: Assume P is invertible. Let $p_1, \ldots, p_n$ be the columns of P.
To show: $\{p_1, \ldots, p_n\}$ is linearly independent.
Assume $c_1, \ldots, c_n \in \mathbb{F}$ and $c_1p_1 + \cdots + c_np_n = 0$.
Let $c = (c_1, \ldots, c_n)^t \in \mathbb{F}^n$.
Since $c_1p_1 + \cdots + c_np_n = 0$ then $Pc = 0$.
So $c = P^{-1}Pc = P^{-1}0 = 0$.
So $c_1 = 0, \ldots, c_n = 0$.
$\Leftarrow$: Assume the columns of P are linearly independent.
To show: There exists $Q \in M_n(\mathbb{F})$ such that $QP = 1$.
Let $p_1, \ldots, p_n$ be the columns of P.
Since $B = \{p_1, \ldots, p_n\}$ is linearly independent and $\dim(\mathbb{F}^n) = n$ then B is a maximal linearly independent set.
Thus, by Theorem 5.3, B is a basis.
Let $S = \{e_1, \ldots, e_n\}$ where e_i has 1 in the ith spot and 0 elsewhere.
Then $P = P_{BS}$, the change of basis matrix from S to B.
Let $Q = P_{SB}$, the change of basis matrix from B to S.
Then $QP = P_{SB}P_{BS} = P_{SS} = 1$.
So P is invertible. □

7.4 Some proofs: Inner product spaces

7.4.1 Expansion in an orthonormal basis

Proposition 7.26. *Assume $B = (b_1, \ldots, b_n)$ is an ordered orthonormal basis of V and $x \in V$. Then*

$$x = \langle x, b_1\rangle b_1 + \cdots + \langle x, b_n\rangle b_n.$$

Proof. Let $x \in V$ and let $c_1, \ldots, c_n \in \mathbb{F}$ be such that $x = c_1b_1 + \cdots + c_nb_n$.
If $j \in \{1, \ldots, n\}$ then

$$\begin{aligned}\langle x, b_j\rangle &= \langle c_1b_1 + \cdots + c_nb_n, b_j\rangle = c_1\langle b_1, b_n\rangle + \cdots + c_n\langle b_n, b_j\rangle \\ &= c_1 \cdot 0 + \cdots + c_{j-1} \cdot 0 + c_j \cdot 1 + c_{j+1} \cdot 0 + \cdots + c_n \cdot 0 = c_j \cdot 1 = c_j.\end{aligned}$$

So $x = \langle x, b_1\rangle b_1 + \cdots + \langle x, b_n\rangle b_n$. □

7.4.2 Gram-Schmidt

Theorem 7.27. *(Gram-Schmidt) Let V be an $\mathbb{F}$-vector space with an inner product $\langle, \rangle \colon V \times V \to \mathbb{F}$.*
(a) Let $\{p_1, p_2, \ldots, p_n\}$ be a basis of V. Define $b_1 = p_1$ and

$$b_{k+1} = p_{k+1} - \frac{\langle p_{k+1}, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \cdots - \frac{\langle p_{k+1}, b_k\rangle}{\langle b_k, b_k\rangle} b_k, \qquad \textit{for } k \in \mathbb{Z}_{>0}.$$

Then $(b_1, b_2, \ldots, b_n)$ is an orthogonal basis in V.
(b) Let $(b_1, \ldots, b_n)$ be an orthogonal basis of V. Define

$$u_1 = \frac{b_1}{\|b_1\|}, \quad \ldots\ , \quad u_n = \frac{b_n}{\|b_n\|}.$$

Then $(u_1, \ldots, u_n)$ is an orthonormal basis of V.

Proof. (Sketch) The proof is by induction on n.
For the base case, there is only one vector b_1 and so there is nothing to show.
Induction step: Assume $(b_1, \ldots, b_n)$ are orthogonal.
Let $j \in \{1, \ldots, n\}$. Then

$$\begin{aligned}\langle b_{n+1}, b_j\rangle &= \left\langle p_{n+1} - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle} b_n, b_j \right\rangle \\ &= \langle p_{n+1}, b_j\rangle - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle}\langle b_1, b_j\rangle - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle}\langle b_n, b_j\rangle \\ &= \langle p_{n+1}, b_j\rangle - \frac{\langle p_{n+1}, b_j\rangle}{\langle b_j, b_j\rangle}\langle b_j, b_j\rangle = \langle p_{n+1}, b_j\rangle - \langle p_{n+1}, b_j\rangle = 0 \quad \text{and} \\ \langle b_j, b_{n+1}\rangle &= \left\langle b_j, p_{n+1} - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle} b_n \right\rangle \\ &= \langle b_j, p_{n+1}\rangle - \frac{\overline{\langle p_{n+1}, b_1\rangle}}{\langle b_1, b_1\rangle}\langle b_j, b_1\rangle - \cdots - \frac{\overline{\langle p_{n+1}, b_n\rangle}}{\langle b_n, b_n\rangle}\langle b_j, b_n\rangle \\ &= \langle b_j, p_{n+1}\rangle - \frac{\overline{\langle p_{n+1}, b_j\rangle}}{\langle b_j, b_j\rangle}\langle b_j, b_j\rangle = \langle b_j, p_{n+1}\rangle - \overline{\langle p_{n+1}, b_j\rangle} = 0,\end{aligned}$$

where the identity $\overline{\langle b_k, b_k\rangle} = \langle b_k, b_k\rangle$ and the last equality follow from the assumption that $\langle, \rangle$ is Hermitian. So $(b_1, \ldots, b_{n+1})$ are orthogonal. □

References

[Boualg] N. Bourbaki. *Éléments de mathématique Algèbres*, ?????????

[Mor] A.O. Morris, *Linear algebra, An introduction*, Chapman and Hall, 1982 ISBN:0-412-38100-1

[SS] E. Shult and D. Surowski, *Algebra, A teaching and source book*, Springer 2015.

[St67] R. Steinberg, *Lectures on Chevalley groups*, Notes prepared by John Faulkner and Robert Wilson. Revised and corrected edition of the 1968 original [MR0466335], University Lecture Series 66 American Mathematical Society, Providence, RI, 2016. ISBN: 978-1-4704-3105-1.

[Tay] D.E. Taylor, *The geometry of the classical groups*, Sigma Series in Pure Mathematics, Volume 9, Heldermann Verlag, 1992.

[Wil] J.H. Wilkinson, *The algebraic eigenvalue problem*, Monographs on Numerical Analysis, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1988. ISBN: 0-19-853418-3, MR0950175