Arutomatisierung, Regelungtechnik

Kristóf Cserpes

October 3, 2021

Contents

Chapter 1

Linear Maps

1.1 Vector spaces

A vector space over a field F is a set V that is closed under vector addition $(+)$ and scalar multiplication $(\cdot)$. A vecotr space must fulfill following axioms:

$$\begin{aligned} v, w, u &\in V \\ \alpha, \beta &\in F \\ v + w &= u \\ \alpha \cdot v &= u \\ (\alpha \cdot \beta) \cdot v &= \alpha \cdot (\beta \cdot v) \\ \alpha(v + w) &= \alpha \cdot v + \alpha \cdot w \\ (\alpha + \beta) \cdot w &= \alpha \cdot v + \beta \cdot w \\ 1 \cdot v &= w \end{aligned} \tag{1.1}$$

The vector addition $(V, +)$ form a commutative group.

$$\begin{aligned} (v + w) + u &= v + (w + z) && \text{Associativity} \\ v + 0 &= V && \text{Identity element: zero vector} \\ v + w &= 0 && \text{Inverse element} \\ v + w &= w + v && \text{Commutativity} \end{aligned}$$

$(F, +, \cdot)$ form a field.

$$\begin{aligned} a, a^{-1}, b, c &\in F \\ (a + b) + c &= a + (b + c) && \text{Additive associativity} \\ a + a^{-1} &= 0 && \text{Additive inverse} \\ a + 0 &= a && \text{Additive identity} \\ a + b &= b + a && \text{Additive commutativity} \\ (a \cdot b) \cdot c &= a \cdot (b \cdot c) && \text{Mulitlicative associativity} \\ a \cdot a^{-1} &= 1, a^{-1} \neq 0 && \text{Mulitlicative inverse} \\ a \cdot 1 &= a && \text{Multiplicative identity} \\ a \cdot b &= b \cdot a && \text{Mulitplicative commutativity} \\ a \cdot (b + c) &= a \cdot b + \cdot c && \text{Distibutivity} \end{aligned}$$

Example for fields: $(\mathbb{R}, +, \cdot)$ $(\mathbb{C}, +, \cdot)$
Example for non fields: $(\mathbb{N}, +, \cdot)$ $(\mathbb{Z}, +, \cdot)$

1.1.1 Subspace

Let V be a vector space ofer a field F. A subspace W is a subset of V that also form a vector space over F.

Example 1:

$\mathbb{R}^2$ is a vector space over $\mathbb{R}$.

$$W = \{c \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \mid c \in \mathbb{R}\} = \{\dots \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2.4 \\ 4.8 \end{bmatrix} \dots\}$$

Direct sum

Let V be a vector space ofer a field F and W and U subspaces of V When:

$$W \cap U = \{0\} \tag{1.2}$$
$$W \cup U = V \tag{1.3}$$

Then U and W are a direct sum of V. It is denoted by

$$V = U \oplus W \tag{1.4}$$

1.1.2 Span, (lineare Hülle)

Let V be a vector space ofer a field F and S a finite subset of V wiht length n. The span of S is the set of vectors that can be created by linear combinations with the vectors in S.

$$span(S) = \{\sum_{i=1}^{n} a_i \cdot s_i \mid n \in \mathbb{N}, a_i \in F, s_i \in S\} \tag{1.5}$$

Example 1:

$$S = \{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\}$$
$$\operatorname{span} S = \{a_1 \begin{bmatrix}1\\0\\0\end{bmatrix} + a_2 \begin{bmatrix}0\\1\\0\end{bmatrix} \mid a_1, a_2 \in \mathbb{R}\}$$

Spanning set

Let V be a vector space ofer a field F and S a finite subset of V. S is a spanning set of if

$$\operatorname{span} S = V \tag{1.6}$$

Example 2:

Let V be $\mathbb{R}^2$ over the field $\mathbb{R}$. Following subsets are spanning set of V:

$$S_1 = \{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$$
$$S_2 = \{\begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}0\\2\end{bmatrix}\}$$
$$S_3 = \{\begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}0\\2\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}\}$$

1.1.3 Base

Let V be a vector space ofer a field F and B a spanning set of V. If the elements of B are lineary independent then B is called a basis. The coefficients of the linear combination are referred to as components or coordinates of the vector with respect to B. The elements of B are called basis vectors.

Example 1:

Let V be $\mathbb{R}^2$ over the field $\mathbb{R}$. Following subsets are spanning set of V:

$$B_1 = \{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$$
$$B_2 = \{\begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}0\\2\end{bmatrix}\}$$

In previous example S_3 is not a valid base, because its elements are lineary dependent.

Standard base

A base B is called a Standard base if the vecotors of B are all zero, except one that equals 1. The vectors of the standard base are called unit vectors.

Example 2:

The standard base for $\mathbb{R}^n$ is

$$\hat{i} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \hat{j} = \begin{bmatrix}0\\1\\0\end{bmatrix}, \hat{k} = \begin{bmatrix}0\\0\\1\end{bmatrix}$$
$$B = \{\hat{i}, \hat{j}, \hat{k}\}$$

A vector v expressed in the standard basis B.

$$v = \begin{bmatrix}4\\5\\6\end{bmatrix} = 4\hat{i} + 5\hat{j} + 6\hat{k}$$

1.1.4 Dimension

The dimension dim of a vector space is the size of its base B. The dimension is equal to the rank (see 2.1.2) of the tranformation matrix.

Example 1:

The vector space $\mathbb{R}^n$

$$\dim \mathbb{R}^n = n$$

Example 2:

$$V = \{c_1 \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, c_2 \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \mid c_1, c_2 \in \mathbb{R}\}$$
$$\dim V = 2$$

Example 3:

$$V = \{c_1 \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, c_2 \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \mid c_1, c_2 \in \mathbb{R}\}$$
$$\dim V = 1$$

Example 4:

The only vector space with dimension 0 is where V contains only the zero vector.

$$\dim\{\begin{bmatrix}0\\0\\0\end{bmatrix}\} = 0$$

sectionLinear maps Let V W be vector spaces over the same field F. A function $f : V \to W$ is said to be a linear map if for any two vectors $v, u \in V$ and any scalar $c \in F$ the following two conditions are satisfied:

$$f(u+v) = f(u) + f(v) \tag{Additvity}$$
$$f(c \cdot u) = c \cdot u \tag{Homogenity}$$

1.1.5 Transformation matrix

Each linear transformation can be represented as a matrix vector multiplication.

$$f : W \to V$$
$$\dim W = n, \dim V = m$$
$$f(x) = A^{m \times n} x$$

Example 1:

$$f(x) = \begin{bmatrix}1 & 2 & 3\\5 & 7 & 9\end{bmatrix} \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$
$$f\left(\begin{bmatrix}5\\6\\7\end{bmatrix}\right) = \begin{bmatrix}38\\130\end{bmatrix}$$

Composition

If there are two linear maps f, g wiht tranformation matrices:

$$f : V \to W = Ax$$
$$g : U \to V = Bx$$

then the compostition is:

$$h : U \to W = f \circ g$$
$$h(x) = A(Bx) = (A \cdot B)x$$

Example 2:

$$
\begin{aligned}
f(x) &= \begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 9 \end{bmatrix} \cdot x \\
g(x) &= \begin{bmatrix} -1 & -2 \\ -7 & -9 \\ 13 & 17 \end{bmatrix} \cdot x \\
h(x) = f \circ g &= \begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 9 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ -7 & -9 \\ 13 & 17 \end{bmatrix} \cdot x = \begin{bmatrix} 24 & 31 \\ 64 & 80 \end{bmatrix} \cdot x \\
g\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) &= \begin{bmatrix} -11 \\ -57 \\ 107 \end{bmatrix}, f\left(\begin{bmatrix} -11 \\ -57 \\ 107 \end{bmatrix}\right) = \begin{bmatrix} 196 \\ 509 \end{bmatrix} \\
h(\begin{bmatrix} 3 \\ 4 \end{bmatrix}) &= \begin{bmatrix} 196 \\ 509 \end{bmatrix}
\end{aligned}
$$

1.1.6 Image (Bild)

The image $f^{\rightarrow}$ of a tranformation $L : V \rightarrow W$ is the set of vectors that the tranformation can produce.

$$f^{\rightarrow}(L) = \{L(x) \mid x \in V\} \tag{1.7}$$

The image is the columnspan of the tranformation matrix. The dimenson of the image ist called **rank**, and is the same as the rank of the transformation matrix.

Example 1:

$$
\begin{aligned}
L(x) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x \\
f^{\rightarrow}(L) &= \mathbb{R}^2
\end{aligned}
$$

Example 2:

$$
\begin{aligned}
L(x) &= \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} x \\
f^{\rightarrow}(L) &= \{ \begin{bmatrix} c_1 \\ 2c_1 \\ c_1 + c_2 \\ 2c_2 \end{bmatrix} \mid c_1, c_2 \in \mathbb{R}\}
\end{aligned}
$$

1.1.7 Kernel, Null Space (Kern)

The kernel of a linear map $L : V \rightarrow W$ is the linear subspace of the domain of the map which is mapped to the zero vector.

$$\ker L = \{v \in V | L(v) = 0\} \tag{1.8}$$

The vecots of the kernel are the set of vectors that yield the zero vector after multiplication with the tranformation matrx.

$$
\begin{aligned}
&L : Ax = y \\
&x' \in \ker L \textbf{ if } Ax' = 0
\end{aligned}
$$

The kernel forms a subspace of V:

$$
\begin{aligned}
v, u \in \ker L, \alpha \in \mathbb{F} \\
\alpha v \in \ker L \\
v + u \in \ker L
\end{aligned}
$$

The dimension of the kernel is called the **nullity**.

Example 1:

$$L : Ax = y$$
$$A = \begin{bmatrix} 1 & -1, & 0 \\ 0 & -2 & 4 \end{bmatrix}$$

To calculate $\ker A$ simply set y to the zero vector and solve for x.

$$\begin{bmatrix} 1 & -1, & 0 \\ 0 & -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$x_1 - x_2 = 0 \to x_1 = x_2$$
$$-2x_2 + 4x_3 = 0 \to x_2 = 2x_3$$

The kernel is:

$$\ker L = \{c \cdot \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \mid c \in \mathbb{C}\}$$

A concrete example:

$$\begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix} \in \ker L$$
$$L\left(\begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 \cdot 4 - 1 \cdot 4 + 0 \cdot 2 \\ 0 \cdot 4 - 2 \cdot 4 + 4 \cdot 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Example 2:

1.2 Rank-Nullity theorem

If $L : V \to W$ is a linear tranformation then it.

$$\operatorname{rank} L + \operatorname{nullity} T = \dim(imageT) + \dim(\ker(T)) = \dim(V) \tag{1.9}$$

Example 1:

$$L : \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x = y$$
$$imageT = \mathbb{R}^2$$
$$\operatorname{rank} L = 2$$
$$\ker T = \{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\}$$
$$\operatorname{nullity} L = 0$$
$$\dim V = \operatorname{rank} L + \operatorname{nullity} T = 2 + 0 = 2$$

Example 2:

$$L : \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix} x = y$$
$$imagL = \{\begin{bmatrix} c \\ 2c \end{bmatrix} \mid c \in \mathbb{R}\}$$

$$\operatorname{rank} L = 1$$
$$\ker L = \{\begin{bmatrix} -2c_1 - 4c_2 \\ c_1 \\ c_2 \end{bmatrix} \mid c_1, c_2 \in \mathbb{R}\} \operatorname{nullity} A = 2$$

Chapter 2

Matrices

2.1 Properties

2.1.1 Dimension

The dimension [1] is the number of rows a and columns b of a Matrix A

$$\dim A = a \times b \tag{2.1}$$

Denoted as:

$$A^{a \times b}$$

Example 1:

$$\dim \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = 2 \times 3$$

Example 2:

Linearly dependent rows/columns

$$\dim \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = 2 \times 2$$

Matlab

`size` (API link)

```
A = [[1,2,3],[1,2,3]]
size(A)
ans  2  3
```

2.1.2 Rank (Rang)

Rowsapce, columnspace

The rowspace C of a matrix ist the span of its column vectors.
The definied as is the span of its row vectors. It is dentoed as $C(A^T)$
The dimension of the column and rowspace are always equal.

Example 1:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 & 4 \end{bmatrix}$$
$$C(A) = \{ \begin{bmatrix} c \\ c \end{bmatrix} \mid c \in \mathbb{R} \}$$
$$C(A^T) = \{ \begin{bmatrix} c \\ 2c \\ 4c \end{bmatrix} \mid c \in \mathbb{R} \}$$
$$\dim C(A) = \dim C(A^T) = 1$$

[1] Not to be confused with the dimenson of a vector space, see 1.1.4

Rank

The rank of a matrix A is the maximal number of linearly independent columns (or the number of linearly independent rows, is the same thing). Or equally, the rank of a matrix A is the dimenson of its columnspace (or rowspace):

$$\operatorname{rank} A = \dim C(A) = \dim C(A^T) \tag{2.2}$$

Example 2:

$$\operatorname{rank} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = 2$$

Example 3:

Both rows are linearly dependent

$$\operatorname{rank} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} = 1$$

Example 4:

Only a matrix containing zeroes has a rank of 0

$$\operatorname{rank} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0$$

Example 5:

Both columns are linearly independent, some rows are linearly dependent.

$$\operatorname{rank} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 5 & 7 \end{bmatrix} = 2$$

Matlab

`rank` (API link)

```
A = [[1,2,3],[1,2,3]]
rank(A)
ans = 1
```

2.1.3 Trace (Spur)

The trace of a square matrix A is the sum of all its main diagonal elements.

$$tr(A) = \sum_{i=0}^{n} a_{ii} \tag{2.3}$$

Example 1:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$
$$tr(A) = 1 + 5 + 9 = 15$$

Matlab

`trace` (API link)

```
>> A = [1,2,3;4,5,6;7,8,9]
>> trace(A)
ans =
15
```

2.1.4 Minor, Cofactors

Submatrix

A submatrix S_ij of a Matrix A is the Matrix obtained by deleting the ith Row and deleting the jth column.

Example 1:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$S_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}$$

Minor

A minor M_{ij} of a matrix A is the determinant of the submatrix S_{ij}.

Cofactors

A cofactor C_{ij} is obtained by multiplying the minor M_{ij} by $(-1)^{i+j}$. The cofactor Matrix C is given by:

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1i} \\ C_{21} & C_{22} & \dots & C_{1i} \\ \vdots & \vdots & \ddots & \\ C_{j1} & C_{j2} & & C_{ij} \end{bmatrix} = \begin{bmatrix} M_{11} & -M_{12} & \dots & (-1)^{i+1}M_{1i} \\ -M_{21} & M_{22} & \dots & (-1)^{i+2}M_{2i} \\ \vdots & \vdots & \ddots & \\ (-1)^{1+j}M_{j1} & (-1)^{2+j}M_{j2} & & (-1)^{i+j}M_{ij} \end{bmatrix} \tag{2.4}$$

Example 2:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$M_{11} = det(\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}) = -3 \qquad M_{12} = det(\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}) = -6 \qquad M_{13} = det(\begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}) = -3$$

$$M_{21} = det(\begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix}) = -6 \qquad M_{22} = det(\begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}) = -12 \qquad M_{23} = det(\begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}) = -6$$

$$M_{31} = det(\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}) = -3 \qquad M_{32} = det(\begin{bmatrix} 1 & 4 \\ 3 & 6 \end{bmatrix}) = -6 \qquad M_{33} = det(\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}) = -3$$

$$C = \begin{bmatrix} M_{11} & -M_{12} & M_{13} \\ -M_{21} & M_{22} & -M_{23} \\ M_{31} & -M_{32} & M_{33} \end{bmatrix} = \begin{bmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{bmatrix}$$

Example 3:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$M_{11} = 4 \qquad M_{12} = 3$$

$$M_{21} = 2 \qquad M_{22} = 1$$

$$C = \begin{bmatrix} M_{11} & -M_{12} \\ -M_{21} & M_22 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}$$

2.1.5 Determinant

2x2 Matrix

For 2x2 Matrix the formula is as given:

$$\det\left(\begin{bmatrix} x_1, x_2 \\ x_3, x_4 \end{bmatrix}\right) = x_1 \cdot x_4 - x_2 \cdot x_4$$

Example 1:

$$\det\left(\begin{bmatrix} 3 & 7 \\ -5 & 11 \end{bmatrix}\right) = 3 \cdot 11 - 7 \cdot (-5) = 68$$

3x3 Matrix

The determinant of a 3x3 Matrix can be calculated using its minors.

$$\begin{aligned}\det\left(\begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9) \end{bmatrix}\right) &= x_1 \cdot \det\left(\begin{bmatrix} x_5 & x_6 \\ x_8 & x_9 \end{bmatrix}\right) - x_2 \cdot \det\left(\begin{bmatrix} x_4 & x_6 \\ x_7 & x_9 \end{bmatrix}\right) + x_3 \cdot \det\left(\begin{bmatrix} x_4 & x_5 \\ x_7 & x_8 \end{bmatrix}\right) \\ &= x_1(x_5x_9 - x_6x_8) - x_2(x_4x_9 - x_6x_7) + x_3(x_4x_8 - x_5x_7) \\ &= x_1x_5x_9 + x_2x_6x_7 + x_3x_4x_8 - x_3x_5x_7 - x_2x_4x_9 - x_1x_6x_8\end{aligned}$$

For higher order matrices you can apply this method recursively.

Example 2:

Minors were calculated in previous example.

$$\det\left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\right) = 1 \cdot (-3) - 2 \cdot (-6) + 3 \cdot (-3) = 0$$

Triangular matrx

$$D = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ 0 & x_{22} & & \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & x_{nn} \end{bmatrix} \det(D) = x_{11} \cdot x_{22} \ldots x_{nn} = \prod_{i=1}^{n} x$$

Singular matrix

Singular matrices are matrices with det = 0. Singular matrices have rows and/or columns that are not linearly independent.

Example 3:

$$A = \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}$$
$$\det(A) = 1 \cdot (-4) - (-2) \cdot 2 = 0$$

Matlab

`det` (API link)

```
>> a =[[3,7];[4,12]]
>> det(a)
ans = 8
```

2.1.6 Eigenvalues, Eigenvectors

An eigenvector v of a square matrix A is a nonzero vector that changes at most by a scalar factor when that linear transformation is applied to it. The corresponding eigenvalue λ is the factor by which the eigenvector is scaled.

$$A \cdot v = \lambda \cdot v \tag{2.5}$$

Example 1:

$$A = \begin{bmatrix} 5 & 1 \\ 0 & 3 \end{bmatrix}$$
$$v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \lambda = 5$$
$$A \cdot v = \begin{bmatrix} 5 & 1 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

Characteristic polynomial

The expression 2.5 can be written as:

$$A \cdot v = \lambda \cdot I \cdot v \qquad \text{Multiplying with identity Matrix}$$
$$A \cdot v - \lambda \cdot I \cdot v = 0 \tag{2.6}$$
$$v \cdot (A - \lambda \cdot I) = 0 \tag{2.7}$$

Since v per definition can't be the zero vector, the expression $(A - \lambda \cdot I)$ must be zero.

$$A - \lambda \cdot I = 0$$
$$\det(A - \lambda \cdot I) = 0$$
$$\det\left(\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22}-\lambda & & \\ \vdots & & \ddots & \\ a_{n1} & & & a_{nn}-\lambda \end{bmatrix}\right) = 0$$

The characteristic polynomial P_A of a matrix A is defined as:

$$P_A(t) = \det(A - tI) \tag{2.8}$$

If a square matrix A with $\dim(A) = n \times n$ then $p_A(t)$ will have a degree of n. The characteristic polynomial is always monic (the leading coefficient is 1)

Example 2:

$$A = \begin{bmatrix} 5 & 7 \\ 11 & 3 \end{bmatrix}$$
$$p_A(t) = \det\left(\begin{bmatrix} 5-t & 7 \\ 11 & 3-t \end{bmatrix}\right) = (5-t)\cdot(3-t) - 7 \cdot 11 = t^2 - 8t - 62$$

Note for a 2×2 matrix $p_A(t)$ is always:

$$p_A(t) = t^2 - \text{trace}(A)t - \det(A) \tag{2.9}$$

Matlab

`charpoly` (API link)

```
>> charpoly([5, 7 ; 11, 3])
ans =
     1     -8     -62
```

If A gets pluged into $p_A(t)$ then the result will be the zero-matrix.

$$P_a(A) = A^n + b_2 A^{n-1} \cdots b_{n-1}A + b_n I = 0 \tag{2.10}$$

Example 3:

From previous example.

$$p_A(t) = t^2 - 8t - 62$$
$$p_A(A) = \begin{bmatrix} 5 & 7 \\ 11 & 3 \end{bmatrix}^2 - \begin{bmatrix} 5 & 7 \\ 11 & 3 \end{bmatrix} - 62I$$
$$= \begin{bmatrix} 102 & 56 \\ 88 & 86 \end{bmatrix} - \begin{bmatrix} 40 & 56 \\ 88 & 24 \end{bmatrix} - \begin{bmatrix} 62 & 0 \\ 0 & 62 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Characteristic equation

The roots of the characteristic polynomial are the eigenvalues λ_i of A. The expression

$$p_A(t) = 0 \tag{2.11}$$

is called the characteristic equation. The characteristic polynomial can be written as:

$$p_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_i) \tag{2.12}$$

Example 4:

$$A = \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$$

$$p_A(t) = \det(A - tI) = \det\left(\begin{bmatrix} 3-t & 7 \\ 2 & 5-t \end{bmatrix}\right) = t^2 - 8t + 1$$

We get the eigenvalues by setting $p_A(\lambda) = 0$ and solving for λ

$$\lambda^2 - 8\lambda + 1 = 0$$
$$\lambda_{12} = -\frac{-8}{2} \pm \sqrt{\left(\frac{-8}{2}\right)^2 - 1}$$
$$\lambda_1 = 4 - \sqrt{15}, \lambda_2 = 4 + \sqrt{15}$$

Arithmetic Multiplity

A matrix can have multiple eigenvalues λ_i with the same value. The characteristic polynomial can be written as:

$$p_A(t) = (t - \lambda_1)(t - \lambda_2)\dots(t - \lambda_n)$$

The arithmetic Multiplicity $\mu_A(\lambda_1)$ is the number of times $(t - \lambda_i)$ can divide $p_A(t)$, so the highest power $(t - \lambda_i)$ can have (simply said the number of times a value appears).

Example 5:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

A has 4 eigenvalues: 1, 2, 3, 4($= \lambda_{1..10}$)
The characteristic polynomial can be expressed by using only distinct eigenvalues:

$$p_A(t) = (t-1)(t-2)^2(t-3)^3(t-4)^4$$

For example $\mu_A(\lambda_4) = 4$, because $(t-4)$ divides $p_A(t)$ 4 times.

Eigenvectors, eigenspace

To find the eigenvector of an associatited eigenvalue we need to find the kernel of the following linear map:

$$L : (A - \lambda_i \cdot I)x = y$$
$$\epsilon_i = \ker L$$

Since the kernel of a tranformation forms a vectorspace ϵ called **eigenspace**. So following properites are satisfied:

$$v_1, v_2 \in \epsilon_i, c \in \mathbb{F}$$
$$v_1 + v_2 \in \epsilon_i$$
$$c \cdot v_1 \in \epsilon_i$$

Example 6:

Continuing example 2.

$$A = \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$$

$$\lambda_1 = 4 - \sqrt{15}, \lambda_2 = 4 + \sqrt{15}$$

λ_1:

$$\begin{bmatrix} 3-(4-\sqrt{15}) & 7 \\ 2 & 5-(4-\sqrt{15}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} -1+\sqrt{15} & 7 \\ 2 & 1+\sqrt{15} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We can eliminate the II row by subtracting I $\left(\frac{2}{-1+\sqrt{15}}\right)$

$$\begin{bmatrix} -1+\sqrt{15} & 7 \\ 2 & 1+\sqrt{15} \end{bmatrix} \rightarrow \begin{bmatrix} -1+\sqrt{15} & 7 \\ 2-2 & (1+\sqrt{15})-\left(\frac{14}{-1+\sqrt{15}}\right) \end{bmatrix} = \begin{bmatrix} -1+\sqrt{15} & 7 \\ 0 & 0 \end{bmatrix}$$

Since the last row was eliminated, we see that of $rank(A-\lambda I)$ is 1. It means x_1 or x_2 can be freely chosen.

Keep in mind we are interest only in the 'form' of the eigenvector, because an eigenvector of A multiplied with a scalar is still an eigenvector of A.

$$0 = (-1+\sqrt{15})x + 7y$$
$$y = \frac{(1-\sqrt{15})x}{7}$$

We can eliminate the fraction by setting $x = 7$.

$$x = 7$$
$$y = \frac{(1-\sqrt{15})7}{7} = 1-\sqrt{15}$$
$$v_1 = \begin{bmatrix} 7 \\ 1-\sqrt{15} \end{bmatrix}$$

Same for λ_2:

$$\begin{bmatrix} 3-(4+\sqrt{15}) & 7 \\ 2 & 5-(4-\sqrt{15}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1-\sqrt{15} & 7 \\ 2 & 1-\sqrt{15} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We can eliminate the II row by subtracting I $\left(\frac{2}{-1-\sqrt{15}}\right)$

$$\begin{bmatrix} -1-\sqrt{15} & 7 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving for x_1, x_2:

$$(-1-\sqrt{15})x_1 + 7x_2 = 0$$
$$x_2 = \frac{1+\sqrt{15}x_1}{7}$$
$$x_1 = 7 \text{ (chosen)}$$
$$x_2 = 1+\sqrt{15}$$
$$v_2 = \begin{bmatrix} 7 \\ 1+\sqrt{15} \end{bmatrix}$$

Matlab

`eig` (API link)

```
    >> [a, d] = eig([3, 7; 11, 3])
a =
     0.6236     -0.6236
     0.7817      0.7817
d =
     11.7750           0
     0      -5.7750
```

Geometric multiplicity

The geometry multiplicity γ_A of an eigenvalue is the dimension of the associatited eigenspace.

$$\gamma_a(\lambda_i) = \dim \ker (A - I\lambda_i) \tag{2.13}$$

The geometry multiplicity of an eigenvalue can't be larger than the arithmetic multiplicity.

$$\gamma_A(\lambda_i) \leq \mu_a(\lambda_1) \tag{2.14}$$

For a square Matrix $A^{n\times n}$ with m eigenvalues it holds:

$$\sum_{i=1}^{m} \gamma(\lambda_i) + \mu(\lambda_i) = n \tag{2.15}$$

Example 7:

$$\begin{aligned}
A = &\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\
p_A(t) &= t^3(t-2) \\
\lambda_1 &= 2, \lambda_2 = 0 \\
\mu_A(2) = 1, \mu_A(0) &= 3, \\
\epsilon_1 = &\{ \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \} \\
\epsilon_2 = &\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \} \\
\gamma_a(2) &= \dim \epsilon_1 = 1 \\
\gamma_a(0) &= \dim \epsilon_2 = 2
\end{aligned}$$

2.1.7 Similarity

Two square matrices A and B are similar when an if there exists an invertible $n \times n$ matrix P such that:

$$A = P^{-1}BP \tag{2.16}$$
$$B = PAP^{-1} \tag{2.17}$$

It is denoted as

$$A \tilde{=} B$$

U is also called the change of base matrix. Similar matrices have the same:

- Characteristic polynomial
- Eigenvalues (but not eigenvectors)
- Determinant
- Trace

Similarity is an equivalence relation

- A is similar to A
- If A is similar to B, then B is similar to A.
- If A is similar to B and B is similar to C, then A is similar to C.

Example 1:

$$B = \begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix}, A = \begin{bmatrix} 3 & 4 \\ \frac{1}{4} & 3 \end{bmatrix}, P = \begin{bmatrix} 3 & 0 \\ 1 & 4 \end{bmatrix}, P^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ -\frac{1}{12} & \frac{1}{4} \end{bmatrix}$$

$$P^{-1}BP = \begin{bmatrix} \frac{1}{3} & 0 \\ -\frac{1}{12} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 1 \\ -\frac{1}{6} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ \frac{1}{4} & 3 \end{bmatrix}$$

$$\begin{aligned}
\det(B) &= 2 \cdot 4 - 3 \cdot 0 = 8 \\
\det(A) &= 3 \cdot 3 - 1 \cdot \frac{1}{4} = 8 \\
tr(A) &= 3 + 3 = 6 \\
tr(B) &= 2 + 4 = 6 \\
p_B(t) &= (2-t)(4-t) - 4 \cdot 0 = t^2 - 6t + 8 \\
p_A(t) &= (3-t)(3-t) - 4 \cdot \frac{1}{4} = t^2 - 6t + 8
\end{aligned}$$

2.1.8 Defective matrices

If there is one eigenvalue λ_i with $\mu_A(\lambda_i) \neq \gamma_A(\lambda_i)$ then the corresponding Matrix A defective:

- The matrix has less than n lienary independent eigenvectors
- The sum of the dimesons of the eigensapces has a dimension less than n

The eigenvalue is called defective eigenvalue.

Example 1:

$$\begin{gathered}
A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix} \\
\lambda_1 = 1, \lambda_2 = 4 \\
v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} \frac{1}{9} \\ \frac{1}{3} \\ 1 \end{bmatrix} \\
\mu_A(\lambda_1) = 2 \neq \gamma_A(\lambda_1) = 1
\end{gathered}$$

Defective matrices can't be diagonalized.

2.1.9 Geeralized eigenvectors

Let $L : V \to V$ with a defective transformation matrix A. A generalized eigenvector w of a defective eigenvalue is the solution of:

$$\begin{aligned}
(A - \lambda I)^m w &= 0 \\
(A - \lambda I)^{m-1} w &\neq 0 \\
m > 1, m &\in \mathbb{N}
\end{aligned} \tag{2.18}$$

m is called the rank of the generalized eigenvector.

Jordan chain

Let v be an ordinary eigenvector of A:

$$\begin{aligned}
(A - \lambda I)v &= 0 \\
(A - \lambda I)w_1 &= v \\
(A - \lambda I)w_2 &= w_1 \\
(A - \lambda I)w_3 &= w_2 \\
&\vdots \\
(A - \lambda I)w_{n-1} &= w_n
\end{aligned}$$

Example 1:

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

The matrix has single eigenvalue $\lambda = 3$ with $\mu_a(3) = 3$ but only one eigenvector v_1.

$$A - I\lambda = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix} v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix} w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, w_1 = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

2.1.10 Shift matrix

A shift matrix that has 1 on its superdiagonal and 0 elsewhere.

$$S = \begin{bmatrix} 0 & 1 & 0 & \cdots & & \\ \vdots & 0 & 1 & 0 & \cdots \\ \vdots & & \ddots & & \\ & & & 0 & 1 \\ & & \ddots & 0 & 0 \end{bmatrix}$$

When multplied wiht another matrix A it shifts the columns of A by one to the right.

$$A = \begin{bmatrix} c_1 & c_2 & \cdots & c_{n-1} & c_n \end{bmatrix}$$
$$A \cdot S = \begin{bmatrix} 0 & c_2 & \cdots & c_{n-2} & c_{n-1} \end{bmatrix}$$

Example 1:

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

2.1.11 Nilpotent matrix

A matrix is nilpotent of degree k if

$$A^i \neq 0$$
$$A^k = 0$$
$$0 \leq i < k$$

Example 1:

Let A be a Matrix containing only zeroers except on its superdiagonal. A is nilpotent of degree $k + 1$ where k is the number of nonzero element. An example would be the shift matrix:

$$S^1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} S^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} S^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Example 2:

Only the zero matrix is nilpotent with degree 1.

Example 3:

A diagonal matrix is not nilpotent.

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$
$$D^n = \begin{bmatrix} 2^n & 0 \\ 0 & 4^n \end{bmatrix}$$

2.1.12 Jordan normal form

Jordan Block

A jordan block is a square matrix with the same value for each element on its main diagonal and 1 on it superdiagonal. The other elements are 0.

$$B_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & \\ 0 & \lambda & 1 & 0 & \cdots \\ \vdots & & \ddots & & \\ & & & \lambda & 1 \\ & & \ddots & 0 & \lambda \end{bmatrix}$$

Jordan box

Let *lambda* be an eigenvalue of A with $\mu(\lambda) = n$ and $\gamma_a(\lambda) = m$. A Jordan box is the direct sum of

$$J_\lambda = D_\lambda \oplus B_\lambda \tag{2.19}$$

where:

$$\dim J_\lambda = n$$
$$\dim D_\lambda = m$$
$$\dim B_\lambda = n - m$$

Example 1:

$$\mu(\lambda) = 4, \gamma(\lambda) = 2$$
$$D_\lambda = \begin{bmatrix} \lambda \end{bmatrix} \; B_\lambda = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$
$$J_\lambda = \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

Example 2:

$$\mu(\lambda) = 4, \gamma(\lambda) = 3$$
$$D_\lambda = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \; B_\lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$
$$J_\lambda = \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

2.2 Operations

2.2.1 Transposing

Transpose of a matrix A is an operator which flips a matrix over its diagonal; that is, it switches the row and column indices of the matrix A by producing another matrix, often denoted by A^T.

Example 1:

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Example 2:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Notice the diagonal elements do not get swaped by transposing. So for any diagonal matrix D holds $D = D^T$.

Matlab

`transpose` (API link)

```
A = [1,2,3;4,5,6]
transpose(A)
ans =
1     4
2     5
3     6
```

2.2.2 Direct sum

The direct sum of two matrixes $A^{a\times b}$ and $B^{c\times d}$ is defined as

$$\begin{aligned} C = A \oplus B &= \begin{bmatrix} A & N_1 \\ N_2 & B \end{bmatrix} \\ \dim C &= (a+c) \times (b+d) \end{aligned} \tag{2.20}$$

N_1 and N_2 are zero matrices with dimensons $\dim N_1 = b$

2.2.3 Diagonalisation

A matrix A is diagonalizabe if A is similar (see 2.1.7) to a diagonal matrix D.

$$D = U^{-1}AU$$

Eigendecomposition

A matrix can be diagonalized using its eigenvalues and eigenvectors. D is a diagonal matrix containing the eigenvalues λ_i on it main diagonal. The the eigenspaces ϵ_i form a base called **eigenbase** (when the arithmetic multiplicity of an eigenvalue is 1 then ϵ is just the eigenvector). So change of base matrix U has the base vectors of the eigenspaces as it's columns.

$$A = UDU^{-1} \tag{2.21}$$

$$U = \begin{bmatrix} v_1 & v_2 \cdots v_n \end{bmatrix} \tag{2.22}$$

Example 1:

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 5 & 0 & -1 \end{bmatrix}$$

The eigenvalues λ_i and eigenvectors v_i are:

$$\lambda_1 = 4, v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda_2 = -2, v_2 = \begin{bmatrix} -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda_3 = 2, v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

We can construct U and D. Keep in mind that the order of the eigenvalues in the diagonal of D must match the order of the order of eigenvector columns in U (and U^{-1}).

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

$$U = \begin{bmatrix} -\frac{1}{5} & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$U^{-1} = \begin{bmatrix} \frac{5}{6} & 0 & \frac{1}{6} \\ -\frac{5}{6} & 0 & \frac{5}{6} \\ 0 & 1 & 0 \end{bmatrix}$$

The diagonalized A is:

$$A = \begin{bmatrix} -\frac{1}{5} & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} \frac{5}{6} & 0 & \frac{1}{6} \\ -\frac{5}{6} & 0 & \frac{5}{6} \\ 0 & 1 & 0 \end{bmatrix}$$

Matlab

```
>> a = [3,0,1;0,2,0;5,0,-1]
>> [v, d] = eig(a)
   v =
       0.7071    -0.1961         0
            0          0    1.0000
       0.7071     0.9806         0
   d =
       4     0     0
       0    -2     0
       0     0     2
>> v*d*inv(v)
       ans =
             3.0000         0    1.0000
                  0    2.0000         0
             5.0000         0   -1.0000
```

2.2.4 Raising a matrix to the nth power using diagonalisation

Using the definition of matrix multiplication a single squaring a matrix takes $O(n^3)$ computation steps. Raising a matrix to the mth power would take $O(m \cdot n^3)$ steps. For any diagonal matrix it holds:

$$D^m = \begin{bmatrix} x_{11}^m & & & \\ & x_{22}^m & & \\ & & \ddots & \\ & & & x_{nn}^m \end{bmatrix} \tag{2.23}$$

Using diagonalisation a more effcient calculation can be achieved:

$$\begin{aligned} A &= UDU^{-1} \\ A^2 = A \cdot A &= UD\cancel{U^{-1}U}DU^{-1} = UD^2U^{-1} \\ A^3 = A^2 \cdot A &= UD^2\cancel{U^{-1}U}DU^{-1} = UD^3U^{-1} \\ &\vdots \\ A^n &= UD^nU^{-1} \end{aligned}$$

2.2.5 Matrix exponential

The taylor series of the exponential functions is given as:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} \cdots \tag{2.24}$$

Using this definition you can define the matrix exponential:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = I + A + \frac{A^2}{2} + \frac{A^3}{6} \cdots \tag{2.25}$$

Diagonal Case

$$\begin{aligned} e^D &= I + \begin{bmatrix} x_{11} & & & \\ & x_{22} & & \\ & & \ddots & \\ & & & x_{nn} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} x_{11}^2 & & & \\ & x_{22}^2 & & \\ & & \ddots & \\ & & & x_{nn}^2 \end{bmatrix} + \frac{1}{6}\begin{bmatrix} x_{11}^3 & & & \\ & x_{22}^3 & & \\ & & \ddots & \\ & & & x_{nn}^3 \end{bmatrix} + \frac{1}{24}\begin{bmatrix} x_{11}^4 & & & \\ & x_{22}^4 & & \\ & & \ddots & \\ & & & x_{nn}^4 \end{bmatrix} + \cdots \\ &= I + \begin{bmatrix} x_{11} & & & \\ & x_{22} & & \\ & & \ddots & \\ & & & x_{nn} \end{bmatrix} + \begin{bmatrix} \frac{x_{11}^2}{2} & & & \\ & \frac{x_{22}^2}{2} & & \\ & & \ddots & \\ & & & \frac{x_{nn}^2}{2} \end{bmatrix} + \begin{bmatrix} \frac{x_{11}^3}{6} & & & \\ & \frac{x_{22}^3}{6} & & \\ & & \ddots & \\ & & & \frac{x_{nn}^3}{6} \end{bmatrix} + \begin{bmatrix} \frac{x_{11}^4}{24} & & & \\ & \frac{x_{22}^4}{24} & & \\ & & \ddots & \\ & & & \frac{x_{nn}^4}{24} \end{bmatrix} + \cdots \\ &= \begin{bmatrix} \sum_{m=0}^{\infty} \frac{x_{11}^m}{m!} & & & \\ & \sum_{m=0}^{\infty} \frac{x_{22}^m}{m!} & & \\ & & \ddots & \\ & & & \sum_{m=0}^{\infty} \frac{x_{22}^m}{m!} \end{bmatrix} = \begin{bmatrix} e^{x_{11}} & & & \\ & e^{x_{22}} & & \\ & & \ddots & \\ & & & e^{x_{nn}} \end{bmatrix} \end{aligned}$$

2.2.6 Diagonalizable case

If A is diagonalizabe with UDU^{-1} then:

$$\begin{aligned}
e^A &= \sum_{n=0}^{\infty} \frac{UD^nU^{-1}}{n!} = UIU^{-1} + \frac{UDU^{-1}}{n!} + \frac{UD^2U^{-1}}{2} + \frac{UD^3U^{-1}}{6} \cdots \\
&= U^{-1}\left(\sum_{n=0}^{\infty} \frac{D^n}{n!}\right) U = U^{-1}e^D U \\
&= U^{-1} \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix} U
\end{aligned}$$

Example 1:

$$A = \begin{bmatrix} 3 & -4 \\ -5 & -5 \end{bmatrix}$$

$$\lambda_1 = -7, v_1 = \begin{bmatrix} \frac{2}{5} \\ 1 \end{bmatrix}$$

$$\lambda_2 = 5, v_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$e^D = \begin{bmatrix} \frac{2}{5} & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-7} & 0 \\ 0 & e^5 \end{bmatrix} \begin{bmatrix} \frac{5}{12} & \frac{5}{6} \\ \frac{-5}{12} & \frac{1}{6} \end{bmatrix}$$

Matlab

`expm` (API link)

```
>> a = [3, -4; -5, -5]
>> expm(a)
   ans =
        123.6778   -49.4707
        -61.8384    24.7363
```

Comuting matrices

If two matrices A and B commute ($AB = BA$) then

$$e^{A+B} = e^A e^B \tag{2.26}$$

Jordan bock

A Jordan block B_λ of size m can be separated into:

$$B_\lambda = D_\lambda + S \tag{2.27}$$

where S is the shift matrix (the shift matrix is nilpotent of degree m).

$$e^{B_\lambda} = e^{D_\lambda + S} = e^{D_\lambda} \cdot e^S \qquad \text{(Note } D_\lambda \text{ and S commute)}$$

$$\begin{aligned}
&= e^{D_\lambda} \cdot e^S = e^{D_\lambda}\left(\sum_{n=0}^{m-1} \frac{S^n}{n!} = I + S + \frac{S^2}{2} + \frac{S^3}{6} \cdots\right) \\
&= e^{D_\lambda}\left(I + \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & & & & \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & & & & \end{bmatrix} + \cdots + \frac{1}{(m-1)!}\begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & & 0 \\ \vdots & & & \end{bmatrix}\right) \\
&= e^{D_\lambda} \begin{bmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{6} & \cdots & \frac{1}{(m-1)!} \\ 0 & 1 & 1 & \frac{1}{2} & \cdots & \frac{1}{(m-2)!} \\ 0 & 0 & 1 & 1 & \cdots & \frac{1}{(m-3)!} \\ \vdots & & & & & \end{bmatrix} = \begin{bmatrix} e^\lambda & e^\lambda & \frac{1}{2}e^\lambda & \frac{1}{6}e^\lambda & \cdots & \frac{1}{(m-1)!}e^\lambda \\ 0 & e^\lambda & e^\lambda & \frac{1}{2}e^\lambda & \cdots & \frac{1}{(m-2)!}e^\lambda \\ 0 & 0 & e^\lambda & e^\lambda & \cdots & \frac{1}{(m-3)!}e^\lambda \\ \vdots & & & & & \end{bmatrix}
\end{aligned}$$

Example 2:

$$B_2 = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$e^{B_2} = \begin{bmatrix} e^2 & e^2 & \frac{1}{2}e^2 & \frac{1}{6}e^2 \\ 0 & e^2 & e^2 & \frac{1}{2}e^2 \\ 0 & 0 & e^2 & e^2 \\ 0 & 0 & 0 & e^2 \end{bmatrix}$$

Chapter 3

Differential Equation

3.1 Homogeneous linear first-order

The homogeneous linear first-order differential equations have the form:

$$f'(t) + p(t)f(t) = 0$$

Homogeneous is because one side of the equation is zero. You can rewrite the expression above to have $f(x)$ separated

$$f'(t) = -p(t)f(t) \tag{3.1}$$

$$f'(t)\frac{1}{f(t)} = -p(t) \tag{3.2}$$

Now all the $f(t)$ terms are on the left hand side.

Note the following differentiation:

$$(\ln f(t))' = \frac{1}{f(t)}f'(t) \tag{3.3}$$

It can be used to help integrate 3.2.

$$\int f'(t)\frac{1}{f(t)} = \int -p(t)$$
$$\ln(|f(t)|) + C_1 = -P(t) + C_2$$
$$\ln(|f(t)|) = -P(t) + \hat{C}$$

You can combine C_1 and C_2 to a single constant $\hat{C}$, because they both are constants. Since the domain of ln is $(0, \inf]$ you have to take the absolute value of $f(t)$. To get rid of ln raise both side to e. To compensate for the absolute value you have to take $\pm$ of e.

$$|f(t)| = e^{-P(t)+\hat{C}} \tag{3.4}$$

$$f(t) = \pm e^{-P(t)+\hat{C}} \tag{3.5}$$

$$f(t) = e^{-P(t)}C \tag{3.6}$$

The expression $e^{\hat{C}}$ is a constant, so it can be replaced by C, which constant be $\pm$.

The expression 3.6 is the general solution for homogeneous first order linear differential equations. Any linear combination of the general is a valid solution to the differential equation.

Notation

The notation of differential equations can be simplified by:

$$f(t) = y$$
$$f(t)' = y'$$

Example 1:

$$y' + sin(x+2)y = 0$$

The general solution is:

$$p(t) = sin(x+2)$$
$$P(t) = -cos(x+2)$$
$$y = e^{-p(t)}C$$
$$y = e^{cos(x+2)}C$$

Following functions are valid solutions to the homogeneous equation:

$$y = 2e^{cos(x+2)}$$
$$y = 2e^{cos(x+2)} + 3e^{cos(x+2)}$$
$$y = 0$$

Non-homogenous

$$y' + p(t)y = s(t) \tag{3.7}$$

Recall the product rule which states

$$(f \cdot g) = f'g + fg' \tag{3.8}$$

If we multiply 3.7 by an unknown function $\mu(t)$ called the **integrating factor** we get the following expression:

$$\mu(t)y' + \mu(t)p(t)y = \mu(t)s(t) \tag{3.9}$$

we have an expression that looks like the product rule, if we suppose that

$$\mu(t)' = \mu(t)p(t) \tag{3.10}$$
$$\mu(t) = \int \mu(t)p(t) \tag{3.11}$$

Integrating 3.9 by using the product formula in reverse we get:

$$\int \mu(t) \cdot y' + \mu(t) \cdot p(t) \cdot ydt = \int \mu(t) \cdot s(t)dt \tag{3.12}$$
$$\mu(t)y = \int \mu(td)s(t)dt \tag{3.13}$$
$$y = \frac{1}{\mu(t)} \int \mu(t)s(t) \tag{3.14}$$

The function that satisfies 3.10 is

$$\mu(t) = C(t)e^{P(t)} \tag{3.15}$$
$$\mu(t)' = C(t)'e^{P(t)} + C(t) \cdot p(t) \cdot e^{P(t)} \tag{3.16}$$

We can chose $C(t) = 1$

$$\mu(t) = e^{P(t)} \tag{3.17}$$
$$\mu(t)' = p(t)e^{P(t)} = p(t)\mu(t) \tag{3.18}$$
$$\tag{3.19}$$

Setting 3.17 in 3.14 we get the solution for the differential equation.

$$y = e^{-P(t)} \int e^{P(t)}s(t)dt \tag{3.20}$$

In general integrating the expression above yield the following expression

$$y = y_p + y_h \tag{3.21}$$

Where y_h is the solution to the homogenous equation $y' + p(t)y = 0$. The term y_p is called a particular solution and it is one of the solutions to the nonhomogeneous equation.

Example 2:

$$y' + 2y = t$$
$$p(t) = 2$$
$$P(t) = 2t$$
$$s(t) = t$$
$$y = e^{-2t} \int e^{2t}tdt$$
$$y = e^{-2t}\left(\frac{1}{2}e^{2t} \cdot t - \frac{1}{4}e^{2t} + C\right)$$
$$y = \frac{1}{2}t - \frac{1}{4} + Ce^{-2t}$$

We see that Ce^{-2t} is the solution to the homogeneous equation. $y_p = \frac{1}{2}t - \frac{1}{4}$ is one of the solution to the nonhomogeneous equation.

Matlab

`dsolve` (API link)

```
>> syms y(t)
>> eqn = diff(y, t) + 2*y == t
>> dsolve(eqn)
ans =
     t/2 + (C1*exp(-2*t))/4 - 1/4
```

Note: C1 / 4 is still a constant C.

3.2 Linear first-oder system of differential equations

The exponential map

The exponential map is defined by

$$P_n(t) = e^{At} \tag{3.22}$$

where A is a square matrix with dimension $n \times n$ (see also 2.2.5). If a is diagonalizabe then:

$$P_n(t) = Ue^{Dt}U^{-1} \tag{3.23}$$

The derivate is given by:

$$P_n(t)\frac{d}{dt} = Ae^{At} \tag{3.24}$$

System of differential equations

A system off differential equations contains a set of unkown function $x_1(t), x_2(t) \cdots x_n(t)$ denoted by $x_1, x_2 \cdots x_n$. The derivate of a funciton x_i (with respext to to t) is denoted by x_i'.

3.2.1 Homogenous case

A homogenous linear first-order system has the form:

$$\begin{aligned} &x_1' + p_{11}x_1 + p_{12}x_2 \cdots p_{1n}x_{1n} = 0 \\ &x_2' + p_{21}x_1 + p_{22}x_2 \cdots p_{1n}x_{2n} = 0 \\ &\vdots \\ &x_3' + p_{n1}x_1 + p_{n2}x_2 \cdots p_{1n}x_{nn} = 0 \end{aligned}$$

The terms p_{ij} are called coefficient functions (so written out the are like $p_{ij}(t)$) they only depend on t. The derivates can be expressed as a column vector:

$$\boldsymbol{x'} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix}$$

The coefficient functions can be expressed as a square matrix $A^{n\times n}$

$$\boldsymbol{p} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{3.25}$$

The functions can be written as a column vector

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \tag{3.26}$$

The wholse system can be written in matrix form:

$$\boldsymbol{x'} = \boldsymbol{px} \tag{3.27}$$

The solution is given by (same formula as 3.6 applied to matrices):

$$\boldsymbol{x} = M_p(t)\boldsymbol{c} \tag{3.28}$$

where $\boldsymbol{c}$ is a column vector of constants.

Example 1:

$$\begin{aligned} x_1' &= 1x_1 + 4x_2 \\ x_2' &= 3x_2 + 2x_2 \end{aligned}$$

Which can be rewritten as:

$$\boldsymbol{x'} = \boldsymbol{px}$$
$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The eigenvalues and eigenvectors are:

$$\lambda_1 = -2, v_1 = \begin{bmatrix} \frac{-4}{3} \\ 1 \end{bmatrix} \lambda_2 = 5, v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Calculating $M_p(t)$

$$M_p(t) = \begin{bmatrix} \frac{-4}{3} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-2t} & 0 \\ 0 & e^{5t} \end{bmatrix} \frac{1}{7} \begin{bmatrix} -3 & 3 \\ 4 & 4 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 4e^{-2t} + 3e^{5t} & 4e^{5t} - 4e^{-2t} \\ -3e^{-2t} + 3e^{5t} & 3e^{-2t} + 4e^{5t} \end{bmatrix}$$

The solution is (note the factor $\frac{1}{7}$ can be ignored because $\frac{c}{7}$ ist still a constant):

$$\boldsymbol{x} = M_p(t)\boldsymbol{c} \tag{3.29}$$
$$x_1(t) = c_1 \left(4e^{-2t} + 3e^{5t}\right) + c_2 \left(4e^{5t} - 4e^{-2t}\right) \tag{3.30}$$
$$x_2(t) = c_1 \left(-3e^{-2t} + 3e^{5t}\right) + c_2 \left(3e^{-2t} + 4e^{5t}\right) \tag{3.31}$$

The constans can be simplified by setting $c_a = (3c_1 + 4c_2)$ and $c_b = 3(c_2 - c_1)$:

$$\begin{aligned} x_1(t) &= c_1 4e^{-2t} + c_1 3e^{5t} + c_2 4e^{5t} - c_2 4e^{-2t} = -4(c_2 - c_1)e^{-2t} + (3c_1 + 4c_2)e^{5t} \\ &= c_a e^{5t} - \frac{4}{3} c_b e^{-2t} \\ x_2(t) &= c_1(-3)e^{-2t} + c_1 3e^{5t} + c_2 3e^{-2t} + c_2 4e^{5t} = 3(c_2 - c_1)e^{-2t} + (3c_1 + 4c_2)e^{5t} \\ &= c_b e^{-2t} + c_a e^{5t} \end{aligned}$$

Matlab

```
>> syms x1(t) x2(t)
>> ode1 = diff(x1) == 1*x1 + 4*x2
>> ode2 = diff(x2) == 3*x1 + 2 *x2
>> S = dsolve([ode1 ; ode2])
>> S.x1
   ans =
   C1*exp(5*t) - (4*C2*exp(-2*t))/3
>> S.x2
   ans =
   C2*exp(-2*t) + C1*exp(5*t)
```

A more simplified solution is:

$$\boldsymbol{x} = \sum_{i=1}^{n} c_i \boldsymbol{v_i} e^{\lambda_i t} \tag{3.32}$$