MAT2002 Ordinary Differential Equations
System of first-order equations I

Dongdong He

The Chinese University of Hong Kong (Shenzhen)

March 18, 2021

Overview

1. System of first order equations
   - First-order ODE systems
   - Review of matrices

Outline

Motivation: system of first order equations

Up until now, we have been focusing on ordinary differential equations, where there is **one** independent variable and **one** dependent variable. However, many interesting problems involve multiple ordinary differential equations, leading to a system of equations, that is still **one** independent variable, but now **many** dependent variables. Below we list some examples:

Motivation: system of first order equations

Example: SIR model for disease spreading

Suppose there is a disease among the people in a certain region, the disease can transmit from human to human (such as measles, mumps and rubella).
Without considering the birth and death, the SIR model is one of the simplest compartmental models, the model consists of three compartments:

- $S(t)$ denote the number of susceptible people,
- $I(t)$ be the number of infected people,
- $R(t)$ be the number of recovered people.

Motivation: system of first order equations

Example: SIR model for disease spreading

A person can only be in one of the above three states, and must be susceptible, and then infected, and then recovered. It is not possible for infected people become the susceptible, also it is not possible for recovered people become infected or susceptible (assume the recovered people already got the immunity). Let us assume that

- The decreasing rate of the number of susceptible people is proportional to the rate at which the susceptible people and the infected people meet (proportional to the product of S and I), let β be the contact rate. $S'(t) = -\beta IS, \quad I'(t) = \beta IS.$
- The increasing rate of the number of recovered people is proportional to the number of infected people, the proportional constant is called the recovery rate, let γ be the contact rate. $R'(t) = \gamma I, \quad I'(t) = -\gamma I.$

Motivation: system of first order equations

Example: SIR model for disease spreading

Combine the above two facts, we can model this with the following system of ODEs:

$$S' = -\beta IS, \quad I' = \beta IS - \gamma I, \quad R' = \gamma I,$$

Note that the number of susceptible people is always decreasing, and the number of recovered people is always increasing.

If we sum the three equations we see that

$$\frac{d}{dt}(S + I + R) = 0,$$

and so the **total** number of people is conserved.

Motivation: system of first order equations

Example: predator-prey system

The Lotka-Volterra equations, also known as the predator-prey system, are used to describe the dynamics of biological systems where two species interact, one as a predator (for example, foxes) whose number is $x(t)$ and the other as prey (for example, rabbits) whose number is $y(t)$.

Model assumptions:

- Without predation, the prey are assumed to have unlimited foods and the number of prey grows exponentially. $x'(t) = ax$ (a is the growth rate)
- When there is predation, the decreasing rate of the number of prey is assumed to be proportional to the rate at which the predators and the prey meet. $x'(t) = -bxy$, b can be considered as the predation rate.
- Without prey (prey is the food of predator), the number of predators decay exponentially. $y'(t) = -dy$ (d is the death rare).
- With prey, the increasing rate of predator is also proportional to the rate at which the predators and the prey meet. $y'(t) = cxy$, c can be considered as the the growth rate of predator.

Motivation: system of first order equations

Example: Lotka-Volterra equations-continue

Combing the above facts, the populations change through time according to the pair of equations:

$$x'(t) = ax - bxy, \qquad y'(t) = cxy - dy,$$

where $x(t)$ is the number of prey, $y(t)$ is the number of some predator. a and c are the growth rates of species x and y, b can be seen as a predation rate for species x, and d is a loss rate (death rate) for species y.
The interaction between the two species is modelled through the terms proportional to the product xy.

First order ODE systems

The general system of first order equations involving n dependent variables is

$$\begin{aligned} y_1'(t) &= F_1(t, y_1, \ldots, y_n), \\ y_2'(t) &= F_2(t, y_1, \ldots, y_n), \\ &\vdots \\ y_n'(t) &= F_n(t, y_1, \ldots, y_n), \end{aligned}$$

with initial conditions

$$y_1(t_0) = x_1, \quad \ldots, \quad y_n(t_0) = x_n,$$

where $t_0 \in I$, and $x_1, \ldots, x_n \in \mathbb{R}$ are given, and all $F_i(i = 1, \cdots, n)$ are real functions. solution to the above system can be convenient written as a vector $\mathbf{y}(t) = (y_1(t), \ldots, y_n(t))^T$ where for each $t \in I, \mathbf{y}(t)$ lives in the space $\mathbb{R}^n$. By running through $t \in I$, we then trace out a curve in $\mathbb{R}^n$, which we will call the trajectory, or path. The initial condition $\mathbf{x} = (x_1, \ldots, x_n)^T$ determines the starting point of the path.

First order ODE systems

To ensure that the above IVP has exactly one solution, we state the following existence and uniqueness theorem.

Theorem 9.1

(Existence and Uniqueness). *Suppose the functions* $F_1, \ldots, F_n$ *and the partial derivatives* $\frac{\partial F_1}{\partial y_1}, \frac{\partial F_1}{\partial y_2}, \ldots, \frac{\partial F_1}{\partial y_n}, \frac{\partial F_2}{\partial y_1}, \ldots, \frac{\partial F_2}{\partial y_n}, \ldots, \frac{\partial F_n}{\partial y_n}$ *are all continuous in a region* R *defined as*

$$R := [t_0 - a, t_0 + a] \times [x_1 - b, x_1 + b] \times \cdots \times [x_n - b, x_n - b] \subset \mathbb{R}^{n+1}$$

where a, b *are two positive constants. Then, there is exactly one solution* $\mathbf{y}(t)(t \in (t_0 - h, t_0 + h))$ *to the above IVP, where* $h = \min(a, b/M)$ *and* $M = \max_{(t, x_1, \cdots, x_n) \in R} \{|F_1(t, x_1, \cdots, x_n)|, \cdots, F_n(t, x_1, \cdots, x_n)|\}$.

First order ODE systems

Remark 1

(1) *If $n = 1$, then we have $y' = F(t, y)$, leading to the same assumptions as in the case of first order equations.*

(2) *We only have existence and uniqueness for a small interval $(t_0 - h, t_0 + h)$ around t_0.*

(3) *Note that we can express a $n-$th order (nonlinear) equation as a first order system. Indeed, given the ODE*

$$y^{(n)} = F(t, y, y', \dots, y^{(n-1)}),$$

setting
$$z_1 = y, \qquad z_2 = y', \qquad \dots, \qquad z_n = y^{(n-1)},$$

then we can write down

$$\begin{aligned} z_1' &= z_2, \\ z_2' &= z_3, \\ &\vdots \\ z_n' &= y^{(n)} = F(t, z_1, \dots, z_{n-1}). \end{aligned}$$

First order ODE systems

Definition 9.2

If each F_i, $1 \leq i \leq n$, is linear with respect to $y_1, \ldots, y_n$, then we call the system of ODEs a linear system. Otherwise it is a nonlinear system.

We now study linear systems is greater detail. The general linear system of first order ODEs is

$$\begin{aligned}
y_1'(t) &= p_{11}(t)y_1(t) + p_{12}(t)y_2(t) + \cdots + p_{1n}(t)y_n(t) + g_1(t), \\
y_2'(t) &= p_{21}(t)y_1(t) + p_{22}(t)y_2(t) + \cdots + p_{2n}(t)y_n(t) + g_2(t), \\
&\vdots \\
y_n'(t) &= p_{n1}(t)y_1(t) + p_{n2}(t)y_2(t) + \cdots + p_{nn}(t)y_n(t) + g_n(t),
\end{aligned}$$

where $p_{11}(t), \ldots, p_{nn}(t), g_1(t), \ldots, g_n(t)$ are given real functions.

First order ODE systems

Definition 9.3

It is more convenient to introduce the matrix form. Denoting vectors

$$\mathbf{y}(t) = (y_1(t), \ldots, y_n(t))^T, \quad \mathbf{g}(t) = (g_1(t), \ldots, g_n(t))^T,$$

and the matrix

$$\mathbf{P}(t) = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}(t) & p_{n2}(t) & \cdots & p_{nn}(t) \end{pmatrix}$$

the general system can be written as

$$\mathbf{y}'(t) = \mathbf{P}(t)\mathbf{y}(t) + \mathbf{g}(t).$$

First order ODE systems

Definition 9.4

A first order linear system of equations

$$\mathbf{y}'(t) = \mathbf{P}(t)\mathbf{y}(t) + \mathbf{g}(t)$$

is called **homogeneous** if $\mathbf{g}(t) = \mathbf{0}$, i.e., $g_i(t) = 0$ for $1 \le i \le n$. Otherwise it is called non-homogeneous.

First order ODE systems

The above linear first-order ODE system subject to the following initial conditions

$$y_1(t_0) = x_1, \quad \dots, \quad y_n(t_0) = x_n,$$

is the initial value problem (IVP).

Theorem 9.5

(Existence and uniqueness for linear systems). *Let $I = (\alpha, \beta) \subset \mathbb{R}$ be an open interval such that functions $p_{11}(t), \dots, p_{nn}(t), g_1(t), \dots, g_n(t)$ are continuous in I. For $t_0 \in I$ and $x_1, \dots, x_n \in \mathbb{R}$, there is exactly one solution $\mathbf{y} = (y_1(t), \dots, y_n(t))^T$ to the IVP.*

We will not prove this theorem in this course.

In order to study first order linear systems, we make use of the convenient matrix form. Therefore we need to recall some basic properties of matrices.

Review of matrices

A matrix is a **rectangular array** of numbers. We use the notation $\mathbf{A} \in \mathbb{R}^{m\times n}$ to denote a matrix with real entries of size m rows by n columns. If $\mathbf{A}$ is a complex valued matrix we write $\mathbf{A} \in \mathbb{C}^{m\times n}$. Furthermore, we often wirte

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{or} \quad \mathbf{A} = (a_{ij})_{1\le i\le m, 1\le j\le n}.$$

The **transpose** of a matrix $\mathbf{A}$ is denoted as $\mathbf{A}^T$ which is defined as

$$\mathbf{A}^T = (a_{ji})_{1\le j\le n, 1\le i\le m} \in \mathbb{R}^{n\times m}.$$

Review of matrices

For a complex-valued matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, we define its complex conjugate as

$$\overline{\mathbf{A}} = (\overline{a_{ij}})_{1 \le i \le m, 1 \le j \le n}$$

For example

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3+i \\ i & 4 & -7 \end{pmatrix} \text{ with } \mathbf{A}^T \begin{pmatrix} 1 & i \\ 2 & 4 \\ 3+i & -7 \end{pmatrix}, \; \bar{\mathbf{A}} = \begin{pmatrix} 1 & 2 & 3-i \\ -i & 4 & -7 \end{pmatrix}$$

For two matrices of the same size $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ we can define addition and subtraction, as well as scalar multiplication. For products of matrices, we require a matrix $\mathbf{A} \in \mathbb{R}^{m \times p}$ and another $\mathbf{B} \in \mathbb{R}^{p \times n}$, where the number of columns in $\mathbf{A}$ is equal to the number of rows in $\mathbf{A}$. Then

$$\mathbf{AB} = \left(\sum_{j=1}^{p} a_{ij} b_{jk} \right)_{1 \le i \le m, 1 \le k \le n}.$$

The same also holds for complex-valued matrices. Note that in general

$$\boxed{\mathbf{AB} \neq \mathbf{BA}},$$

i.e., product of matrices is <u>**not commutative**</u>.

Review of matrices

In order to solve a linear system of equations, we can use the matrix notation to write

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

as

$$\mathbf{Ax} = \mathbf{b}, \quad \text{where } \mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m.$$

To find the solution (if one exists) we can apply **elementary row operations** to the augmented matrix $(\mathbf{A}|\mathbf{b}) \in \mathbb{R}^{m\times(n+1)}$. Let us briefly recall the elementary row operations applied to a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$. Denoting the ith rows of $\mathbf{A}$ as $r_i \in \mathbb{R}^n$, we have

(1) interchange two rows $r_i \leftrightarrow r_j$;

(2) non-zero scalar multiple of one row $r_i \mapsto \alpha r_i$, $\alpha \neq 0$;

(3) adding a multiple of one row to another $r_i \mapsto r_i + \alpha r_j$, $\alpha \neq 0$.

Review of matrices

Example: Solving linear system

Solve the linear system

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 7, \\ -x_1 + x_2 - 2x_3 &= -5, \\ 2x_1 - x_2 - x_3 &= 4. \end{aligned}$$

Writing

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 3 \\ -1 & 1 & -2 \\ 2 & -1 & -1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 7 \\ -5 \\ 4 \end{pmatrix},$$

we now apply elementary row operations to the augmented matrix

$$(\mathbf{A}|\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & -2 & 3 & 7 \\ -1 & 1 & -2 & -5 \\ 2 & -1 & -1 & 4 \end{array}\right).$$

Review of matrices

Example: Solving linear system

First apply $r_2 \mapsto r_1 + r_2$ and $r_3 \mapsto r_3 - 2r_1$ leads to

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & 7 \\ 0 & -1 & 1 & 2 \\ 0 & 3 & -7 & -10 \end{array}\right).$$

Then apply $r_2 \mapsto -r_2$ leads to

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & 7 \\ 0 & 1 & -1 & -2 \\ 0 & 3 & -7 & -10 \end{array}\right).$$

Then apply $r_1 \mapsto r_1 + 2r_2$ and $r_3 \mapsto r_3 - 3r_2$ leads to

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & -4 & -4 \end{array}\right).$$

Review of matrices

Example: Solving linear system

Then apply $r_3 \mapsto -\frac{1}{4} r_3$ leads to

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right).$$

Then apply $r_1 \mapsto r_1 - r_3$ and $r_2 \mapsto r_3 + r_2$ leads to

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array} \right),$$

and so the solution is

$$x_1 = 2, \qquad x_2 = -1, \qquad x_3 = 1.$$

Review of matrices

For the case $m = n$, i.e., $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbf{A} \in \mathbb{C}^{n \times n}$) we call $\mathbf{A}$ a **square matrix**. A special square matrix is the identity matrix $\mathbf{I}$ defined as

$$I_{kk} = 1, 1 \leq k \leq n, \qquad I_{ij} = 0, 1 \leq i \neq j \leq n.$$

In particular, all the diagonal entries are one and all other entries are zero.

Definition 9.6

We say a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is **invertible** if there is a matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ such that

$$\boxed{\mathbf{AB} = \mathbf{BA} = \mathbf{I}}.$$

In this case we write $\mathbf{B} = \mathbf{A}^{-1}$. Matrices that do not have an inverse are called **singular** or **non-invertible**.

For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ the following statements are equivalent:

(1) $\mathbf{A}$ is invertible;

(2) the determinant $\det(\mathbf{A})$ is non-zero;

(3) The only solution to the problem $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.

Review of matrices

We can also use the elementary row operations to find the inverse of a square matrix. To do this we consider the augmented matrix $(\mathbf{A}|\mathbf{I})$ and transform this into the matrix $(\mathbf{I}|\mathbf{B})$. It turns out that $\mathbf{B}$ will be the inverse of $\mathbf{A}$.

Review of matrices

Example: find matrix inverse

Find the inverse of the matrix

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

Let us write the augmented matrix $(\mathbf{A}|\mathbf{I})$:

$$(\mathbf{A}|\mathbf{I}) = \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 3 & -1 & 2 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right).$$

First apply $r_2 \mapsto r_2 - 3r_1$ and $r_3 \mapsto r_3 - 2r_1$ leads to

$$\left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 2 & 5 & -3 & 1 & 0 \\ 0 & 4 & 5 & -2 & 0 & 1 \end{array}\right).$$

Review of matrices

Example: find matrix inverse

Then apply $r_2 \mapsto r_2/2$ leads to

$$\left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 5/2 & -3/2 & 1/2 & 0 \\ 0 & 4 & 5 & -2 & 0 & 1 \end{array}\right).$$

Then apply $r_3 \mapsto r_3 - 4r_2$ and $r_1 \mapsto r_1 + r_2$ leads to

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3/2 & -1/2 & 1/2 & 0 \\ 0 & 1 & 5/2 & -3/2 & 1/2 & 0 \\ 0 & 0 & -5 & 4 & -2 & 1 \end{array}\right).$$

Then apply $r_3 \mapsto -r_3/5$ leads to

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3/2 & -1/2 & 1/2 & 0 \\ 0 & 1 & 5/2 & -3/2 & 1/2 & 0 \\ 0 & 0 & 1 & -4/5 & 2/5 & -1/5 \end{array}\right).$$

Review of matrices

Example: find matrix inverse

Then apply $r_1 \mapsto r_1 - 3r_3/2$ and $r_2 \mapsto r_2 - -5r_3/2$ leads to

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 7/10 & -1/10 & 3/10 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 & -4/5 & 2/5 & -1/5 \end{array}\right).$$

Hence, the inverse of $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \begin{pmatrix} 7/10 & -1/10 & 3/10 \\ 1/2 & -1/2 & 1/2 \\ -4/5 & 2/5 & -1/5 \end{pmatrix}.$$

Matrix with entries are continuous functions

In our study of first order systems, we will deal with the case where the entries of the matrix $\mathbf{A}$ are functions of the independent variable t, hence we can define a matrix function of t as $\mathbf{A}(t)$ where

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}(t) & a_{m2}(t) & \cdots & a_{mn}(t) \end{pmatrix}.$$

We say that $\mathbf{A}(t)$ is **continuous** if all the entries $a_{11}(t), \ldots, a_{mn}(t)$ are continuous functions of t. Similarly, we say $\mathbf{A}(t)$ is **differentiable** if all its entries are differentiable functions. Then

$$\frac{d}{dt}\mathbf{A}(t) = \begin{pmatrix} a'_{11}(t) & a'_{12}(t) & \cdots & a'_{1n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a'_{m1}(t) & a'_{m2}(t) & \cdots & a'_{mn}(t) \end{pmatrix}.$$

We can also define the (indefinite) integral of $\mathbf{A}(t)$ as

$$\int \mathbf{A}(t)dt = \left(\int a_{ij}(t)dt \right)_{1 \le i \le m, 1 \le j \le n}.$$

Matrix with entries are continuous functions

$$\frac{d(\mathbf{A}(t)\mathbf{B}(t))}{dt} = \frac{d\mathbf{A}(t)}{dt}\mathbf{B}(t) + \mathbf{A}(t)\frac{d\mathbf{B}(t)}{dt}$$

Example 9.7

For

$$\mathbf{A}(t) = \begin{pmatrix} \cos t & \sin t \\ e^t & t \end{pmatrix},$$

we have

$$\frac{d}{dt}\mathbf{A}(t) = \mathbf{A}'(t) = \begin{pmatrix} -\sin t & \cos t \\ e^t & 1 \end{pmatrix}, \qquad \int \mathbf{A}(t)dt = \begin{pmatrix} \sin t & -\cos t \\ e^t & \frac{1}{2}t^2 \end{pmatrix}.$$

Review of matrices

An important concept of matrix theory involves eigenvalues and eigenvectors.

Definition 9.8

For a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, if there is a number r and a **non-zero** vector $\mathbf{x}$ such that

$$\boxed{\mathbf{Ax} = \lambda \mathbf{x}},$$

then we say λ is an **eigenvalue** of $\mathbf{A}$ with **corresponding eigenvector** $\mathbf{x}$.

Therefore we can see that $\mathbf{x}$ is a **non-zero** solution to the problem

$$\boxed{(\mathbf{A} - \lambda \mathbf{I})\mathbf{y} = \mathbf{0}},$$

and so

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0.$$

In particular, using that fact that the determinant of $(\mathbf{A} - \lambda \mathbf{I})$ can be expressed as a polynomial in r of degree n, which we also term as the **characteristic polynomial of A**, we can find the roots of this polynomial to deduce the eigenvalues.

Review of matrices

Definition 9.9

Let r be an eigenvalue of the matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. We define the **algebraic multiplicity** of r as the multiplicity of r as a root of the characteristic polynomial $p_{\mathbf{A}}(x) = \det(\mathbf{A} - x\mathbf{I})$. The **geometric multiplicity** of r is the number of linearly independent eigenvectors associated to r (i.e., the dimension of the eigenspace for r).

Review: Algebraic multiplicity and geometric multiplicity

Theorem 9.10

Let λ be an eigenvalue of $n \times n$ matrix A, then geometric multiplicity $\leq$ algebraic multiplicity for each eigenvalue.

Proof. Suppose that A is real-valued. And λ_0 is any eigenvalue of A, assume that $dim(Null(\lambda_0 I - A)) = k$, then we can pick k orthonormal eigenvectors $\mathbf{v}_1, \cdots, \mathbf{v}_k$ from $Null(\lambda_0 I - A)$, and we can add $\mathbf{v}_{k+1}, \cdots, \mathbf{v}_n \in \mathbb{R}^n$ such that $V = [\mathbf{v}_1 | \cdots | \mathbf{v}_n] = [V_1 | V_2]$ $(V_1 = [\mathbf{v}_1, \cdots, \mathbf{v}_k], V_2 = [\mathbf{v}_{k+1}, \cdots, \mathbf{v}_n])$ is an orthogonal matrix. Now, it is easy to check that

$$V^T A V = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} [V_1 | V_2] = \begin{bmatrix} V_1^T A V_1 & V_1^T A V_2 \\ V_2^T A V_1 & V_2^T A V_2 \end{bmatrix} = \begin{bmatrix} \lambda_0 I_{k \times k} & V_1^T A V_2 \\ O & V_2^T A V_2 \end{bmatrix}$$

Review: Algebraic multiplicity and geometric multiplicity

It follows that

$$\begin{aligned}
\det(A-\lambda I) &= \det(V^T(A-\lambda I)V) = \det(V^T AV - \lambda I) \\
&= \det\begin{pmatrix} (\lambda_0-\lambda)I & (V_1)^T AV_2 \\ 0 & (V_2)^T AV_2 - \lambda I \end{pmatrix} \\
&= (\lambda_0-\lambda)^q \det((V_2)^T AV_2 - \lambda I)
\end{aligned}$$

Here $\det((V_2)^T AV_2 - \lambda I)$ is a polynomial of degree of $n-k$. From the above equation we see that $\det(A-\lambda I)$ has at least k repeated roots for $\lambda = \lambda_0$. Secondly, the complex eigenvalues and eigenvectors could be proved by extending orthogonal matrix into unitary matrix.
The proof is complete.

Review: Eigenvectors belonging to distinct eigenvalues are linearly independent

Theorem 9.11

*(**Eigenvectors belonging to distinct eigenvalues are linearly independent**)*
If $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of the $n \times n$ matrix A, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ are the eigenvectors corresponding to $\lambda_1, \lambda_2, \cdots, \lambda_k$, respectively, then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ are linearly independent.

Example 9.12

Let $A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$.

The characteristic equation is

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -3 \\ 2 & -5-\lambda \end{vmatrix} = (2-\lambda)(-5-\lambda) + 6 = 0.$$

The eigenvalues are $\lambda_1 = 1, \lambda_2 = -4$.

Through some calculations and choose $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ as the eigenvector w.r.t $\lambda_1 = 1$. And choose $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as the eigenvector w.r.t $\lambda_2 = -4$.

Then $\mathbf{x}_1, \mathbf{x}_2$ are linearly independent.

The Case of diagonalizable matrix

Definition 9.13

A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is **diagonalizable** if there is an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that

$$\boxed{\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}}.$$

For linear algebra, if $\mathbf{A} \in \mathbb{R}^{n \times n}$ has n linearly independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with eigenvalues $r_1, \ldots, r_n$, then

$$\mathbf{P} = (\xi_1 | \xi_2 | \ldots | \xi_n), \qquad \mathbf{D} = \begin{pmatrix} r_1 & 0 & \ldots & 0 \\ 0 & r_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & r_n \end{pmatrix},$$

i.e., the columns of $\mathbf{P}$ are the eigenvectors and $\mathbf{D}$ is the diagonal matrix where the entries of the main diagonal are the eigenvalues, ξ_i is the eigenvector corresponding to the eigenvalue r_i $(i = 1, \cdots, n)$.

Diagonalizable matrix

Recall: if A be a square matrix with size n, then A is diagonalizable if and only if A has n linearly independent eigenvectors.

Theorem 9.14

Let A be a square matrix with size n, then A is diagonalizable if and only if the algebraic multiplicity and geometric multiplicity are the same for each eigenvalue.

Proof. Skipped. See Beezer P410.

Remark: **Not every** matrix is diagonalizable.

Review of matrices

Example: eigenvalue and eigenvector

Let

$$\mathbf{A} = \begin{pmatrix} 8 & -9 \\ 4 & -4 \end{pmatrix}.$$

Then

$$\det(\mathbf{A} - r\mathbf{I}) = \begin{vmatrix} 8-r & -9 \\ 4 & -4-r \end{vmatrix} = (8-r)(-4-r)+36 = r^2 - 4r + 4 = (r-2)^2.$$

Hence the eigenvalues are

$$r_1 = r_2 = 2,$$

i.e., the eigenvalue 2 has an algebraic multiplicity of two.

Review of matrices

Example: eigenvalue and eigenvector

To find eigenvectors, we consider non-zeros vectors $\mathbf{x}$ satisfying

$$(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0} \Rightarrow \begin{cases} 8x_1 - 9x_2 = 2x_1 \\ 4x_1 - 4x_2 = 2x_2. \end{cases}$$

Solving the equations implies we have $2x_1 = 3x_2$ and so we can choose

$$x_1 = 1, \qquad x_2 = -2/3 \qquad \Rightarrow \qquad \mathbf{x} = \begin{pmatrix} 1 \\ -2/3 \end{pmatrix}.$$

However, we do not have enough information to deduce another linearly independent vector, thus there is only one eigenvector corresponding to the eigenvalue $r = 2$. Therefore the geometric multiplicity of $r = 2$ is one.

geometric multiplicity $\leq$ algebraic multiplicity

In general, we have

$$1 \leq \text{geo.mult.} \leq \text{alg.mult.}$$

Review of matrices

Example: geometric multiplicity $\leq$ algebraic multiplicity

Let

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

Computing for its eigenvalues, we first solve

$$\det(\mathbf{A} - r\mathbf{I}) = \begin{vmatrix} 1-r & 2 & 1 \\ 1 & -1-r & 1 \\ 2 & 0 & 1-r \end{vmatrix} = -(r-3)(r+1)^2.$$

Hence, we see that $r_1 = 3$ is an eigenvalue of algebraic multiplicity one while $r_2 = r_3 = -1$ is an eigenvalue of algebraic multiplicity two. To find the corresponding eigenvectors, we first compute

$$\mathbf{A} - 3\mathbf{I} = \begin{pmatrix} -2 & 2 & 1 \\ 1 & -4 & 1 \\ 2 & 0 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{pmatrix}$$

where we have used elementary row operations.

Review of matrices

Example: geometric multiplicity $\leq$ algebraic multiplicity

Hence, if we want to find a non-zero vector **x** such that

$$(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{0} \Leftrightarrow \begin{cases} x_1 - x_3 = 0, \\ x_2 - x_3/2 = 0, \end{cases}$$

and we can choose

$$x_1 = 1, \qquad x_2 = 1/2, \qquad x_3 = 1 \qquad \Rightarrow \qquad \mathbf{x} = \begin{pmatrix} 1 \\ 1/2 \\ 1 \end{pmatrix}.$$

Meanwhile, for the other eigenvalue $r_2 = r_3 = -1$, we see that

$$\mathbf{A} + \mathbf{I} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Review of matrices

Example: geometric multiplicity $\leq$ algebraic multiplicity

Hence, for a non-zero vector $\mathbf{y}$ such that

$$(\mathbf{A}+\mathbf{I})\mathbf{y} = \mathbf{0} \Leftrightarrow \begin{cases} y_1 + y_3 = 0, \\ y_2 - y_3/2 = 0, \end{cases}$$

we can choose

$$y_1 = -t, \qquad y_2 = 1/2, \qquad y_3 = 1 \qquad \Rightarrow \qquad \mathbf{y} = \begin{pmatrix} -1 \\ 1/2 \\ 1 \end{pmatrix}.$$

Unfortunately there is no other choice of $\mathbf{y}$, and thus we only have one eigenvector for the eigenvalue $r = -1$. Therefore the geometric multiplicity of $r = -1$ is one.

Review of matrices

Example with complex eigenvalues

Let

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

Computing for the eigenvalues, we solve

$$\det(\mathbf{A} - r\mathbf{I}) = r^2 + 2r + 5 = 0.$$

The quadratic formula gives

$$r_1 = -1 + 2i, \qquad r_2 = -1 - 2i.$$

Since both eigenvalues are distinct, the algebraic multiplicity and hence the geometric multiplicity are one. To find the eigenvectors, consider

$$\mathbf{A} - r_1\mathbf{I} = \begin{pmatrix} -2-2i & -2 \\ 4 & 2-2i \end{pmatrix} \to \begin{pmatrix} 1+i & 1 \\ 2 & 1-i \end{pmatrix} \to \begin{pmatrix} 1+i & 1 \\ 1 & 1/2-1/2i \end{pmatrix} \to \begin{pmatrix} 1+i & 1 \\ 0 & 0 \end{pmatrix}$$

Review of matrices

Example with complex eigenvalues

Hence, if

$$(\mathbf{A} - r_1\mathbf{I})\mathbf{x}_1 = \mathbf{0} \Rightarrow (1+i)x_1 + x_2 = 0,$$

we can choose

$$x_1 = -1, \quad x_2 = 1 + i \quad \Rightarrow \quad \mathbf{x} = \begin{pmatrix} -1 \\ 1+i \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

that is the eigenvector corresponding to a complex eigenvalue is also complex-valued. Note that we do not need to repeat the computations in order to deduce the eigenvector corresponding to r_2, since $\mathbf{A}$ is a real-valued matrix we see that

$$\overline{(\mathbf{A} - r_1\mathbf{I})} = \mathbf{A} - r_2\mathbf{I}.$$

Review of matrices

Example with complex eigenvalues

Hence, if $\mathbf{x}$ is an eigenvector corresponding to r_1, we have

$$\mathbf{0} = \overline{(\mathbf{A} - r_1\mathbf{I})\mathbf{x}} = (\mathbf{A} - r_2\mathbf{I})\bar{\mathbf{x}},$$

and we can take the eigenvector $\mathbf{y}$ corresponding to r_2 as the complex-conjugate of $\mathbf{x}$, i.e.,

$$\mathbf{y} = \begin{pmatrix} -1 \\ 1 - i \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - i \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The takeaway message is that complex eigenvalues and eigenvectors occur in **conjugate pairs**.

Exercise: Find the eigenvalues and eigenvectors of these matrices

$$\mathbf{A} = \begin{pmatrix} 5 & 3 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} -3 & -5 \\ 3/4 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$