Math 310

Fundamental Matrices

→ Read section 7.7

Consider the system of linear equations;

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} \tag{1}$$

A <u>fundamental matrix</u> for (1) is any matrix $\mathbf{\Psi}(t)$ that satisfies

$$\mathbf{\Psi}'(t) = \mathbf{P}(t)\mathbf{\Psi}(t) \tag{2}$$

Note that (2) is a first order differential equation for an unknown matrix, and as such has a unique solution for every initial data $\mathbf{\Psi}(t_o) = \mathbf{B}$ ($\mathbf{B}$ is any given matrix).

One way to form such a matrix is to use any fundamental set of solutions $\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(n)}(t)$ of (1) as it's columns;

$$\mathbf{\Psi}(t) = \begin{pmatrix} x_1^{(1)}(t) & \cdots & x_1^{(n)}(t) \\ \vdots & & \vdots \\ x_n^{(1)}(t) & \cdots & x_n^{(n)}(t) \end{pmatrix}$$

Check that (2) holds for this matrix.

For this fundamental set of solutions, the general solution of (1) is

$$\begin{aligned} \mathbf{x}(t) &= c_1\mathbf{x}_1^{(1)}(t) + \cdots + c_n\mathbf{x}^{(n)}(t) \\ &= \mathbf{\Psi}(t)\mathbf{c}, \quad \mathbf{c} = (c_1, \ldots, c_n) \end{aligned} \tag{3}$$

The same is true for *any* fundamental matrix $\mathbf{\Psi}(t)$; we can write the general solution of (1) as $\mathbf{\Psi}(t)\mathbf{c}$.

Let $\mathbf{x}(t) = \mathbf{\Psi}(t)\mathbf{c}$ be a solution of (1). If the initial condition is $\mathbf{x}(t_o) = \mathbf{x}_o$, then

$$\begin{aligned} \mathbf{\Psi}(t_o)\mathbf{c} &= \mathbf{x}_o \\ &\rightarrow \mathbf{c} = \mathbf{\Psi}^{-1}(t_o)\mathbf{x}_o \\ &\rightarrow \mathbf{x}(t) = \mathbf{\Psi}(t)\mathbf{\Psi}^{-1}(t_o)\mathbf{x}_o \end{aligned} \tag{4}$$

Now we look at a special type of fundamental matrix. We do this by choosing a special initial data matrix $\mathbf{B}$ in (2). Let's look for one that satisfies $\mathbf{\Psi}(t_o) = \mathbf{I}$, the $n \times n$ identity matrix. Then, writing this fundamental matrix as $\mathbf{\Phi}(t)$, since $\mathbf{\Phi}(t_o) = \mathbf{I} = \mathbf{\Phi}^{-1}(t_o)$, $\mathbf{\Phi}(t_o)^{-1}\mathbf{x}_o = \mathbf{x}_o$, we can write (referring to (4));

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{x}_o \tag{5}$$

How to find $\mathbf{\Phi}(t)$? Here's one way. Take any fundamental set $\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(n)}(t)$ with initial values $\mathbf{x}^{(1)}(t_o), \ldots, \mathbf{x}^{(n)}(t_o)$ $= \mathbf{x}_o^{(1)}, \ldots, \mathbf{x}_o^{(n)}$. From this fundamental set we construct another $\mathbf{y}^{(1)}(t), \ldots, \mathbf{y}^{(n)}(t)$ where $\mathbf{y}^{(j)}(t_o) = \mathbf{e}_j$, $1 \le j \le n$ and $\mathbf{e}_j = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ has all zeros except for a 1 in the j^{th} position. To do this we need to solve these systems of equations;

$$\mathbf{y}^{(j)}(t_o) = c_{j1}\mathbf{x}_o^{(1)} + c_{j2}\mathbf{x}_o^{(2)} + \cdots + c_{jn}\mathbf{x}_o^{(n)} = \mathbf{e}_j, \quad 1 \le j \le n$$

for the unknowns c_{jk}. Then we form $\mathbf{\Phi}(t)$ with columns made up of the $\mathbf{y}^{(j)}(t)$.

The exponential of a matrix

In the special case of $\mathbf{x}' = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a constant matrix (so the $\mathbf{P}(t)$ in (1) does not actually depend on time), another way to find $\mathbf{\Phi}(t)$ is as follows. From (2) we see that any fundamental matrix $\mathbf{\Psi}(t)$ of $\mathbf{x}' = \mathbf{A}\,\mathbf{x}$ satisfies

$$\mathbf{\Psi}'(t) = \mathbf{A}\,\mathbf{\Psi}(t) \tag{6}$$

And as before, we will call $\mathbf{\Phi}(t)$ that particular fundamental matrix which satisfies $\mathbf{\Psi}(t_o) = \mathbf{I}$. Recall the scalar case $x' = ax,\ x(t_o) = x_o$ whose solution is $x_o = e^{a(t-t_o)} = x_o \exp(a(t-t_o))$. This suggests we write the solution of (6) as $\mathbf{\Phi}(t) = e^{\mathbf{A}(t-t_o)} = \exp(\mathbf{A}(t-t_o))$ (remember that $\mathbf{\Phi}(t_o) = \mathbf{I}$). How should we define the exponential $\exp(\mathbf{B})$ of a matrix $\mathbf{B}$? Recall the power series;

$$e^{at} \;=\; \exp(at) \;=\; 1 + at + \frac{(at)^2}{2!} + \frac{(at)^3}{3!} + \cdots \;=\; \sum_{i=0}^{\infty} \frac{(at)^i}{i!}, \tag{7}$$

so we may be lead to define $\exp(\mathbf{A}(t-t_o))$ also by a power series;

$$e^{\mathbf{A}(t-t_o)} \;=\; \exp(\mathbf{A}(t-t_o)) \;=\; 1 + \mathbf{A}(t-t_o) + \frac{(\mathbf{A}(t-t_o))^2}{2!} + \frac{(\mathbf{A}(t-t_o))^3}{3!} + \cdots \;=\; \sum_{i=0}^{\infty} \frac{(\mathbf{A}(t-t_o))^i}{i!} \tag{8}$$

It can be shown (along similar lines as done with (7)) that (8) converges for any $\mathbf{A}(t-t_o)$ (any matrix $\mathbf{A}$ and any t, t_o), and so the expression $\exp(\mathbf{A}(t-t_o))$ makes sense as a matrix. Then, following the same lines as for (7), we can differentiate the series (8) and show that

$$\frac{d}{dt}\exp(\mathbf{A}(t-t_o)) \;=\; \mathbf{A}\exp(\mathbf{A}(t-t_o)) \tag{9}$$

Thus, if we define $\mathbf{\Phi}(t)$ by $\mathbf{\Phi}(t) = \exp(\mathbf{A}(t-t_o))$ (as given by formula (8)), then this is the solution of (6) with $\mathbf{\Phi}(t_o) = \mathbf{I}$.

The next task is to actually compute the series (8). We will first look at a simple case when $\mathbf{A} = \mathbf{D}$ is a diagonal matrix with diagonal entries $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then it's easy to see that

$$\mathbf{D}^k \;=\; \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix}$$

(all other entries 0). Therefore,

$$\mathbf{\Phi}(t) \;=\; \exp(\mathbf{D}(t-t_o)) \;=\; \begin{bmatrix} e^{\lambda_1(t-t_o)} & & \\ & \ddots & \\ & & e^{\lambda_n(t-t_o)} \end{bmatrix} \tag{10}$$

(all other entries 0).

The next simplest case is when $\mathbf{A}$ is *diagonalizable.* This means there is an invertible matrix $\mathbf{T}$ such that

$$\mathbf{T}^{-1}\mathbf{A}\,\mathbf{T} \;=\; \mathbf{D} \qquad (\iff \mathbf{A} = \mathbf{T}\,\mathbf{D}\,\mathbf{T}^{-1}) \tag{11}$$

where $\mathbf{D}$ is a diagonal matrix. Recall that $\mathbf{A}$ is diagonalizable if it has n linearly independent eigenvectors $\xi^{(1)}, \ldots, \xi^{(n)}$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues (counted with multiplicity, i.e., if λ occurs as a root of the characteristic equation 2 times, then it appears in the list of eigenvalues 2 times; so not necessarily are all the λ_i distinct); $\mathbf{A}\xi^{(i)} = \lambda_i \xi^{(i)}$.

Given an $\mathbf{x}(t)$ (which we think of as being the solution of $\mathbf{x}' = \mathbf{A}\,\mathbf{x},\ \mathbf{x}(t_o) = \mathbf{x}_o$), let's define the vector $\mathbf{y}(t)$ by

$$\mathbf{x} \;=\; \mathbf{T}\,\mathbf{y} \qquad \Longleftrightarrow \qquad \mathbf{T}^{-1}\mathbf{x} \;=\; \mathbf{y} \tag{12}$$

Then

$$\begin{aligned}
\mathbf{x} \;=\; \mathbf{T}\,\mathbf{y} \Longrightarrow \mathbf{T}\,\mathbf{y}' &= \mathbf{x}' \\
\mathbf{T}\,\mathbf{y}' &= \mathbf{A}\,\mathbf{x} \\
&= \mathbf{A}\,\mathbf{T}\,\mathbf{y} \\
\Longrightarrow \mathbf{y}' &= \mathbf{T}^{-1}\mathbf{A}\,\mathbf{T}\,\mathbf{y} \\
&= \mathbf{D}\,\mathbf{y} \qquad (13)
\end{aligned}$$

So now we can write the solution of (13) as

$$\mathbf{y}(t) \;=\; \exp(\mathbf{D}(t - t_o))\mathbf{y}_o, \quad \mathbf{y}_o = \mathbf{T}^{-1}\mathbf{x}_o \tag{14}$$

where $\exp(\mathbf{D}(t - t_o))$ is given by (10).

We now want to relate $\mathbf{y}(t)$ to the solution $\mathbf{x}(t)$. From (12) and (14),

$$\begin{aligned}
\mathbf{x}(t) &= \mathbf{T}\mathbf{y}(t) \\
&= \mathbf{T}\exp(\mathbf{D}(t - t_o))\mathbf{y}_o \\
&= \mathbf{T}\exp(\mathbf{D}(t - t_o))\mathbf{T}^{-1}\mathbf{x}_o \\
&= \mathbf{\Omega}(t)\mathbf{x}_o, \qquad (15)
\end{aligned}$$

where we've written $\mathbf{\Omega}(t) = \mathbf{T}\exp(\mathbf{D}(t - t_o))\mathbf{T}^{-1}$. Then since $\mathbf{x}'(t) = \mathbf{A}\,\mathbf{x}(t)$, comparing this with (15) implies that $\mathbf{\Omega}\,'(t) = \mathbf{A}\,\mathbf{\Omega}(t)$. Thus, $\mathbf{\Omega}(t)$ is a fundamental matrix of $\mathbf{x}' = \mathbf{A}\,\mathbf{x}$. Note that $\mathbf{\Omega}(t_o) = \mathbf{I}$. Thus,

$$\mathbf{\Phi}(t) \;=\; \mathbf{T}\exp(\mathbf{D}(t - t_o))\mathbf{T}^{-1} \;=\; \mathbf{T}e^{\mathbf{D}(t-t_o)}\mathbf{T}^{-1} \tag{16}$$

is the special fundamental matrix of $\mathbf{x}' = \mathbf{A}\,\mathbf{x}$ with $\mathbf{\Phi}(t_o) = \mathbf{I}$.

Now let $\mathbf{Q}(t) = \exp(\mathbf{D}(t - t_o))$ and $\mathbf{\Psi}(t) = \mathbf{T}\,\mathbf{Q}(t)$. Then

$$\begin{aligned}
\mathbf{\Psi}\,'(t) &= \mathbf{T}\mathbf{Q}\,'(t) \\
&= \mathbf{T}\,\mathbf{D}\,\mathbf{Q}(t) \\
&= \mathbf{T}\,\mathbf{D}(\mathbf{T}^{-1}\mathbf{T}\mathbf{Q} \\
&= (\mathbf{T}\,\mathbf{D}\,\mathbf{T}^{-1})\mathbf{T}\,\mathbf{Q}(t) \\
&= \mathbf{A}\,\mathbf{T}\,\mathbf{Q}(t) \\
&= \mathbf{A}\,\mathbf{\Psi}(t)
\end{aligned}$$

and so $\mathbf{\Psi}(t)$ is a fundamental matrix of $\mathbf{x}' = \mathbf{A}\,\mathbf{x}$ (with $\mathbf{\Psi}(t_o) = \mathbf{T}$).

Note that $\mathbf{\Phi}(t) \;= \mathbf{\Psi}(t)\mathbf{\Psi}^{-1}(t_o)$;

$$\begin{aligned}
\mathbf{\Psi}(t)\mathbf{\Psi}^{-1}(t_o) &= \mathbf{T}\,\mathbf{Q}(t)\Big(\mathbf{T}\,\mathbf{Q}(t_o)\Big)^{-1} \\
&= \mathbf{T}\,\mathbf{Q}(t)\mathbf{Q}^{-1}(t_o)\mathbf{T}^{-1} \\
&= \mathbf{T}\,\mathbf{Q}(t)\,\mathbf{I}\,\mathbf{T}^{-1} \\
&= \mathbf{T}\,\mathbf{Q}(t)\mathbf{T}^{-1} \\
&= \mathbf{\Phi}(t)
\end{aligned}$$

Repeated Eigenvalues

→ Read section 7.8 (and review section 7.3)

A is an $n \times n$ matrix. The characteristic equation $p_A(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$ has roots λ_i which are eigenvalues of A. $p_A(\lambda)$ is a polynomial of order n and so has at most n *real* roots roots (but has exactly n *complex* roots). The associated eigenvectors $\xi^{(i)}$; $\mathbf{A}\xi^{(i)} = \lambda_i\xi^{(i)}$. If eigenvalue λ_i occurs m_i times as a root of $p_A(\lambda)$ we say that λ_i has algebraic multiplicity m_i. Each eigenvalue has at least one associated eigenvector, but it may happen that it has more than one. The geometric multiplicity q_i of the eigenvalue λ_i is the number of linearly independent eigenvectors associated with λ_i. Note that $1 \le q_i \le m_i$. We list the *distinct* eigenvalues as $\lambda_1, \ldots, \lambda_k,\ k \le n,$ and their algebraic multiplicities $m_1, \ldots, m_k$. Note that $m_1 + m_2 + \cdots + m_k = n$.

We now switch to the notation r for the eigenvalues λ, to be consistent with the text.

We consider the system $\mathbf{x}' = \mathbf{A}\,\mathbf{x}$. For each (distinct) eigenvalue r with algebraic multiplicity m we need to construct m linearly independent solutions (then we will have a total of n linearly independent solutions and thus have a fundamental set of solutions).

Case I: Geometric multiplicity $q = m$. Let $\xi^{(1)}, \ldots \xi^{(m)}$ be m linearly independent eigenvectors for r. Then we have m linearly independent solutions

$$e^{rt}\xi^{(1)},\ e^{rt}\xi^{(2)}, \ldots, e^{rt}\xi^{(m)}$$

Case II: $q < m$. Here there are not enough linearly independent eigenvectors to construct independent solutions of the form $e^{rt}\xi$, so we have to look at a way to find more solutions for the eigenvalue r. We will only consider the case when $q = m - 1$, in particular, when $q = 2, m = 1$. Then r occurs twice as a root of $p_A(\lambda)$ but there is only one linearly independent eigenvector ξ. So we have one solution $\mathbf{x}(t) = e^{rt}\xi$. To find another we look for one of the form

$$\mathbf{x}(t) \ = \ \xi t e^{rt} + \eta e^{rt} \qquad (17)$$

for some unknown vector η. Substituting this into the equation $\mathbf{x}' = \mathbf{A}\,\mathbf{x}$ and comparing coefficients of te^{rt} and e^{rt}, we find that

$$\mathbf{A}\,\xi \ = \ r\xi, \quad \text{and} \quad (\mathbf{A} - r\mathbf{I})\eta = \xi \qquad (18)$$

The first equation in (18) is just the eigenvector equation for r (which we know is true), so we need to show that the second equation has a solution, i.e., that there is an η which satisfies that equation. Now, we know that the matrix $\mathbf{A} - r\mathbf{I}$ is not invertible, so we can't solve the equation by $\eta = (\mathbf{A} - r\mathbf{I})^{-1}\xi$. However, the system $(\mathbf{A} - r\mathbf{I})\eta = \xi$ can always be solved for η (say, by Gaussian elimination, row reduction). Thus, we are able to find a second linearly independent solution. The vector η is called a *generalized eigenvector* of $\mathbf{A}$ associated to the eigenvalue r.

It may turn out that when you solve (18) for η, you will get a sum of two or more vectors, one of which may be a multiple of ξ. Since ξ already occurs in the general solution, you can discard this part of η.

See the examples in the text.