APPPHYS217 Tuesday 30 March 2010

Review of Linear Algebra and Ordinary Differential Equations (ODE's) / Part 1 of 2

Matrix and vector notation and operations

We generally think of a vector v as representing a point in R^n. For example,

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad v^T = \begin{bmatrix} x & y & z \end{bmatrix}$$

can be interpreted as a point in 3-space, specified by its Cartesian coordinates. Given two vectors v and w, we can form the "dot product,"

$$v \bullet w = (v^T)w = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = v_1 w_1 + v_2 w_2 + v_3 w_3 = w \bullet v.$$

Taking the square-root of the dot product of a vector with itself gives us its length,

$$|v| = \sqrt{v \bullet v} = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

For example if $v^T = \begin{bmatrix} x & y & z \end{bmatrix}$, then $|v|$ is the familiar expression for the (Euclidean) distance of the point (x,y,z) from the origin.

Note that if we evaluate something like $v(w^T)$ we get a matrix rather than a scalar. With our above definitions of v and w for example,

$$v(w^T) = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} v_1 w_1 & v_1 w_2 & v_1 w_3 \\ v_2 w_1 & v_2 w_2 & v_2 w_3 \\ v_3 w_1 & v_3 w_2 & v_3 w_3 \end{bmatrix},$$

which is in general distinct from

$$w(v^T) = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} v_1 w_1 & v_2 w_1 & v_3 w_1 \\ v_1 w_2 & v_2 w_2 & v_3 w_2 \\ v_1 w_3 & v_2 w_3 & v_3 w_3 \end{bmatrix} = (v(w^T))^T.$$

Just to be clear, the matrix transpose notation means, for example,

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \quad M^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}, \quad (M_1 M_2)^T = M_2^T M_1^T.$$

Note the relevance of the last identity for our vector case: $(v(w^T))^T = w(v^T)$.

In Matlab,

$$v = [1\ 2\ 3]$$

defines a row vector, whereas

$$v = [1;2;3]$$

defines a column vector. Transpose is denoted by $'$ (the apostrophe), such that if

$$v = [1\ 2\ 3];$$
$$w = v';$$

then w is a column vector. Similarly,

$$A = \left[1\ 2\ 3;4\ 5\ 6;7\ 8\ 9\right]$$

is a 3×3 matrix and A' is its transpose.

We can multiply vectors by matrices. Defining

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

we have for example

$$Av = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \\ a_{31}v_1 + a_{32}v_2 + a_{33}v_3 \end{bmatrix},$$

$$(v^T)A = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{21}v_2 + a_{31}v_3 \\ a_{12}v_1 + a_{22}v_2 + a_{32}v_3 \\ a_{13}v_1 + a_{23}v_2 + a_{33}v_3 \end{bmatrix}^T.$$

Here the notation is just meant to emphasize that $(v^T)A$ is a row vector.

Matrices can of course be multiplied and added with each other. If we define

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix},$$

we have

$$AB = \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}) & (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}) & (a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33}) \\ (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}) & (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32}) & (a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33}) \\ (a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31}) & (a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32}) & (a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33}) \end{bmatrix},$$

and in general $AB \neq BA$. One sometimes sees (especially in physics)

$$[A,B] \equiv AB - BA,$$

and if $[A,B] = 0,$ which implies $AB = BA,$ we say that "A and B commute."

In Matlab the $+$ and $*$ symbols are overloaded such that, for example, $1 * 2 = 2$ but $A * B$ is the matrix multiplication of A and $B.$ Likewise for multiplication of vectors by matrices, vector addition, *et cetera*. For example, $\left[1\ 2\ 3\right] * \left[1\ 2\ 3\right]' = 14$ whereas $\left[1\ 2\ 3\right]' * \left[1\ 2\ 3\right]$ is a 3×3 matrix.

Matrices can of course be multiplied by scalars:

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} \\ \alpha a_{21} & \alpha a_{22} & \alpha a_{23} \\ \alpha a_{31} & \alpha a_{32} & \alpha a_{33} \end{bmatrix},$$

where $\alpha \in R$ or $\alpha \in C.$ Note that $\alpha A = (\alpha I)A,$ where

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an identity matrix (here, in R^3). Likewise, one sometimes writes $A + \alpha,$ which usually means

$$A + \alpha \equiv \begin{bmatrix} (a_{11} + \alpha) & a_{12} & a_{13} \\ a_{21} & (a_{22} + \alpha) & a_{23} \\ a_{31} & a_{32} & (a_{33} + \alpha) \end{bmatrix} = A + \alpha I.$$

Note, however, that Matlab will interpret such expressions in a different way! (Exercise: try this for yourself and write an equation to describe what Matlab does.)

Given a (square) matrix $M,$ we can sometimes find a matrix M^{-1} such that

$$M(M^{-1}) = (M^{-1})M = I.$$

We then refer to M^{-1} as the inverse of M (and *vice versa*). For example, if

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad M^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix},$$

as is easily verified by direct multiplication (exercise: prove that the inverse of a matrix, if it exists, is unique). In general for 2×2 matrices,

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad M^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Note that $ad - bc$ here is the determinant of $M.$ For an corresponding expression for 3×3 matrices, see **`http://mathworld.wolfram.com/MatrixInverse.html`** on the same site. For a general discussion of matrix determinants, see **`http://mathworld.wolfram.com/Determinant.html`**. Looking at the general form of an inverse for 2×2 matrices we can already see that M^{-1} won't exist for some matrices $M.$ For example, if

$$M = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix},$$

the determinant is zero and thus our explicit expression for M^{-1} doesn't exist. It is likewise easy to see that if we try to solve

$$M(M^{-1}) = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (a+c) & (b+d) \\ 2(a+c) & 2(b+d) \end{bmatrix} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

we can't simultaneously satisfy $a + c = 1$ and $2(a + c) = 0.$ Hence not all square matrices have inverses. Luckily the invertibility of a matrix is easy to determine: a square matrix M has an inverse if and only if the determinant of M is non-zero. If a matrix has zero determinant we say that it is singular (non-invertible).

In Matlab, the determinant of a matrix can be found using **`det`**, the inverse of a matrix can be found using **`inv`**, and **`cond`** can be used to test whether a matrix is singular (with finite precision arithmetic, **`cond`** provides a more robust method than simply checking **`det`**$(A) = 0$).

Eigenvalues and eigenvectors

Let A be an $n \times n$ matrix with real entries,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad a_{ij} \in R.$$

Then the roots of the equation

$$\det(A - \lambda I) = 0$$

are called the eigenvalues of A. For example, if

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \det(A - \lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 2 \\ 2 & 3-\lambda \end{bmatrix}\right) = (1-\lambda)(3-\lambda) - 4 = 0$$

$$\Rightarrow \lambda^2 - 4\lambda - 1 = 0,$$

whose roots are $\lambda = 2 + \sqrt{5}$ and $\lambda = 2 - \sqrt{5}$. It is important to note that even a matrix with all real entries can have complex (or imaginary) eigenvalues. Consider for example

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

which has $\lambda = 1 \pm i$. One can prove, however, that such complex or imaginary eigenvalues always appear in conjugate pairs: if $\lambda_i \notin R \equiv a + ib$ is an eigenvalue then $\bar{\lambda}_i = a - ib$ is an eigenvalue as well. We'll come back to this important point when we talk about general solutions to linear first-order ODE's.

Sometimes the number of distinct eigenvalues is less than n, as in

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \det(A - \lambda I) = (2-\lambda)(2-\lambda) = 0,$$

which has $\lambda = 2$ as its unique root. As we'll soon see, it makes sense to think of the eigenvalue $\lambda = 2$ in this case as "being repeated twice" or as "having degeneracy two."

Each distinct eigenvalue of a matrix has at least one associated eigenvector. These are defined as the (non-zero) vectors $v_{i,j}$ that satisfy

$$Av_{i,j} = \lambda_i v_{i,j}.$$

In our example immediately above, $\lambda_1 = 2$ has two linearly independent eigenvectors:

$$Av_{1,1} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \lambda_1 v_{1,1},$$
$$Av_{1,2} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \lambda_1 v_{1,2},$$

It is obvious that one cannot write $v_{1,1} = \alpha v_{1,2}$ for any scalar $\alpha,$ which makes it clear that $v_{1,1}$ and $v_{1,2}$ are linearly independent. (One usually assumes that eigenvectors are given in a normalized form; hence if $v_{i,j}$ is presented as an eigenvector of some matrix it is generally safe to assume that $|v_{i,j}| = 1$.) The fact that the eigenvalue $\lambda_1 = 2$ has two independent eigenvectors leads us to think of the eigenvalue as being repeated.

A different sort of situation occurs, for example, for

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad \det(A - \lambda I) = \lambda^2 = 0.$$

This equation has $\lambda_1 = 0$ as its unique root, so this A has only one eigenvalue. However, if we look for solutions of the equation

$$Av_{i,j} = \lambda_i v_{i,j}, \quad (A - \lambda_i I)v_{i,j} = 0,$$
$$\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2b \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

we see that all solutions are going to be linearly dependent. Hence $\lambda_1 = 0$ has only one associated eigenvector in this case and we therefore don't think of the eigenvalue as being repeated.

If an $n \times n$ matrix A has n distinct eigenvalues, or if A has some repeated eigenvalues but is symmetric (by which we mean $A = A^T$), then it is guaranteed that we can choose n linearly-independent eigenvectors for A. As a result, the $n \times n$ matrix P whose columns are these n linearly-independent vectors $v_{i,j}$ is invertible and $P^{-1}AP$ is a diagonal matrix with the eigenvalues of A on the diagonal (ordered in a way that follows from the ordering of the columns of P). Note that eigenvectors corresponding to distinct eigenvalues are necessarily linearly independent. If an $n \times n$ matrix has fewer than n distinct eigenvalues then we need special conditions (such as symmetry) in order to guarantee that we will be able to find n linearly independent eigenvectors, which is in turn what guarantees the existence of P^{-1}.

A rather obvious example is given by

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \lambda_1 = 2, \quad v_{1,1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_{1,2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then $P = P^{-1}$ and $P^{-1}AP = A,$ which is already a diagonal matrix of eigenvalues. Note that any different choice of linearly-independent eigenvectors for the repeated eigenvalue λ_1 works just as well. With the present example, $A = 2I$ so in fact any vector in R^2 is an eigenvector of A corresponding to $\lambda = 2.$ For example,

$$v'_{1,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v'_{1,2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & \sqrt{2} \\ 1 & 0 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 0 & \sqrt{2} \\ 1 & -1 \end{bmatrix},$$

and since $A = 2I$ we can easily verify $P^{-1}AP = 2P^{-1}P = 2I = A.$

For a more typical example, consider

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \quad \lambda_1 = -1, \quad v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \lambda_2 = 2, \quad v_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

for which

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{2}} & \frac{2}{\sqrt{5}} \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} \frac{2\sqrt{2}}{3} & \frac{-\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix},$$

and (incidentally) we see that v_1 and v_2 are not orthogonal:

$$v_1 \cdot v_2 = \frac{-1}{\sqrt{10}}.$$

Still, we recover

$$P^{-1}AP = \begin{bmatrix} \frac{2\sqrt{2}}{3} & \frac{-\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{2}} & \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

This process of bringing a matrix A into diagonal form by a similarity transformation is called diagonalization. It is often useful to think of the matrix P as representing a linear transformation of R^n :

$$v \mapsto v' \equiv P^{-1}v, \quad v, v' \in R^n.$$

For example, next time we will see that one often deals with linear dynamical systems

$$\frac{d}{dt} v = Av,$$

which under the linear transformation P maps to

$$\begin{aligned} \frac{d}{dt}(P^{-1}v) &= (P^{-1}AP)(P^{-1}v) \\ &= D(P^{-1}v), \end{aligned}$$

where D is a diagonal matrix of the eigenvalues of A. We can think of this use of P as "working the problem in different coordinates" such that the system of differential equations become completely decoupled and thus trivial to solve (more on this later). Note that in this context we want P to be invertible so that we can change back to the original coordinate system at the end of our computation.

It is very important to remember that not all matrices are diagonalizable since, as discussed above, we need special conditions to guarantee the existence of an invertible P. In this context, symmetric matrices $A = A^T$ lead to especially simple computations because the eigenvectors of a symmetric matrix can be chosen such that P is orthogonal ($P^{-1} = P^T$). All sorts of nice properties follow from this, which are often exploited in finite-dimensional quantum mechanics calculations. Physicists (such as your Instructor) often make the mistake of assuming that one or more of these nice properties is actually true in general (*i.e.*, for non-symmetric matrices).

In Matlab, eigenvalues and eigenvectors can be found using the function **`eig`**.

Jordan forms

For completeness we include here a very brief discussion of non-diagonalizable matrices.

Consider

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

which has a unique eigenvalue $\lambda_1 = 3,$ and this eigenvalue has only one eigenvector

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Hence we cannot bring A into diagonal form by a similarity transformation (in other words, there is no invertible linear transformation of R^2 that would result in A having diagonal form). In a sense this particular matrix A is already in its simplest possible form (among all forms one could reach via various similarity transforms), called its Jordan Canonical Form. Generally speaking, this is a form that has eigenvalues along the diagonal, ones and zeros on the first super-diagonal, and zeros everywhere else (this includes diagonal matrices). For any matrix $A,$ an invertible linear transformation T can be found that brings A into some Jordan Canonical Form $J,$

$$T^{-1}AT = J.$$

For 2×2 and 3×3 matrices, the complete set of Jordan Canonical Forms is

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix},$$

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

Note that we do not consider row permutations, *et cetera* to be distinct Jordan forms.

One of the reasons that it is can be useful to know the Jordan Canonical Form J of a matrix A is that J is the sum of a diagonal matrix and a nilpotent matrix. A "nilpotent matrix of order r" is a matrix N such that $N^r = 0.$ For example,

$$J = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix} = \lambda_1 I + N, \quad N = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

where

$$N^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$N^3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

This algebraic property can simplify computations involving $J.$ For example for this particular

J, we have

$$J^2 = (\lambda_1 I + N)^2 = \lambda_1^2 I + 2\lambda_1 N + N^2,$$
$$J^3 = (\lambda_1 I + N)(\lambda_1^2 I + 2\lambda_1 N + N^2) = \lambda_1^3 I + 3\lambda_1^2 N + 3\lambda_1 N^2,$$

and for higher powers of J we'll never need to keep more than three powers of N.

Functions of matrices

One often encounters functions of matrices, such as the matrix exponential. These are defined by power series expansion. For example,

$$\exp(A) \equiv \lim_{n\to\infty} \sum_{j=0}^{n} \left(\frac{1}{j!} A^j\right) = 1 + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots.$$

Suppose A is diagonalizable and we have found a linear transformation P of the type discussed above, such that $P^{-1}AP = D$ where D is a diagonal matrix of eigenvalues of A. Then

$$\begin{aligned}
\exp(A) &= \exp(PDP^{-1}) \\
&= 1 + PDP^{-1} + \frac{1}{2}(PDP^{-1})(PDP^{-1}) + \frac{1}{3!}(PDP^{-1})(PDP^{-1})(PDP^{-1}) + \cdots \\
&= 1 + PDP^{-1} + \frac{1}{2}(PD^2P^{-1}) + \frac{1}{3!}(PD^3P^{-1}) + \cdots \\
&= P\left(1 + D + \frac{1}{2}D^2 + \frac{1}{3!}D^3 + \cdots\right)P^{-1} \\
&= P\exp(D)P^{-1},
\end{aligned}$$

where $\exp(D)$ is trivial to compute because it is just the diagonal matrix whose entries are the scalar exponentials of the eigenvalues of A (make sure you can see this for yourself). Even if A is not diagonalizable, we can still simplify the computation of its exponential (to some degree) if we know a linear transformation T that brings it into a Jordan Canonical Form (exercise: try computing $\exp(\alpha J)$ for $\alpha \in R$ and J the 3×3 matrix we considered at the end of the brief section on Jordan forms).

In Matlab, one can compute the matrix exponential using the function **expm**(A). Note that **exp**(A) is also defined but returns the matrix whose entries are the scalar exponentials of the entries of A (exercise: convince yourself that this is a different matrix for general choice of A). Similarly Matlab has functions such as **logm** and **sqrtm** (exercise: show how to compute $A^{1/2}$ for a diagonalizable matrix A), as well as a general-purpose function-to-matrix-function converter **funm**.