Matrix Exponentials, Diagonalization, and Solving Linear Systems

Introduction

In our text we have seen general solutions to 2 dimensional linear systems in the case of two distinct real or two distinct complex eigenvalues or one repeated eigenvalue. Although this provides an easy way to discuss qualitative solutions, it is incomplete in two respects. First, except for the case of a repeated eigenvalue, it does not give us a clean solution to initial value problems. Second, it obscures a commonality that unites all three cases. In this outline I will show how matrix methods reveals this unity, and also extends directly to linear systems with more than two dimensions. The basic approach is described below in outline form. The derivation of the steps in the outline will be provided at the end of the handout.

Outline

1. We consider the system $\mathbf{Y}'(t) = A\mathbf{Y}(t)$ where A is a constant 2×2 matrix with real entries. Let $\mathbf{Y}_0$ be a fixed initial point. The existence and uniqueness theorem says there is a unique solution to the linear system that also satisfies $\mathbf{Y}(t) = \mathbf{Y}_0$ when $t = 0$. We refer to this as the solution of the initial value problem IVP.

2. We define the following matrix function, that is, a 2×2 matrix whose entries are functions of t:

$$e^{At} = I + (At) + \frac{(At)^2}{2} + \frac{(At)^3}{3\cdot 2} + \frac{(At)^4}{4\cdot 3\cdot 2} + \cdots$$

3. A solution of IVP is given by $\mathbf{Y}(t) = e^{At}\,\mathbf{Y}_0$
4. How do we compute e^{At} ? Two cases.
5. Case 1. 2 distinct eigenvalues, either real or imaginary.
    a. Denote the eigenvalues λ and μ, with eigenvectors **v** and **w.**
    b. Form the matrix P whose columns are **v** and **w**.
    c. P has an inverse matrix because eigenvectors for distinct eigenvalues are linearly independent.
    d. A is diagonalizable, and in fact $A = PDP^{-1}$ where $D = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$.
    e. This leads to $e^{At} = Pe^{Dt}P^{-1} = P\begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{bmatrix}P^{-1}$
    f. The solution to the IVP is thus $P\begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{bmatrix}P^{-1} \cdot \mathbf{Y}_0$.

6. Case 2. One (repeated) real eigenvalue λ

   a. In this case the matrix *A* cannot be diagonalized, unless it is already in diagonal form
   b. However, *A* can be *triangularized.* That is, we can find an invertible matrix *P* such that $A = PJP^{-1}$ where $J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$.
   c. This leads to $e^{At} = Pe^{Jt}P^{-1} = P\begin{bmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{bmatrix}P^{-1} = e^{\lambda t}P\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}P^{-1}$
   d. The solution to the IVP is thus $e^{\lambda t}P\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}P^{-1} \cdot \mathbf{Y}_0$.

7. These methods extend to systems with $n \times n$ matrices. If the matrix is diagonalizable then the case 1 development extends in the most obvious way. If the matrix is not diagonalizable, then it can be triangularized to a particular simple form called Jordan Canonical form J. This matrix always has eigenvalues on the diagonal, 1's in certain positions of the superdiagonal (first line of entries above the diagonal), and the exponential matrix e^{Jt} can be worked out fairly simply.

Examples

Example 1. On page 283 of the text, an example is worked out with $A = \begin{bmatrix} 8 & -11 \\ 6 & -9 \end{bmatrix}$. The eigenvalues are found to be $\lambda = -3$ and $\mu = 2$ with corresponding eigenvectors $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 11 \\ 6 \end{bmatrix}$. We put these into the matrix $P = \begin{bmatrix} 1 & 11 \\ 1 & 6 \end{bmatrix}$ and find $P^{-1} = \frac{1}{5}\begin{bmatrix} -6 & 11 \\ 1 & -1 \end{bmatrix}$. Also, our diagonal matrix is $D = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$, so $e^{Dt} = \begin{bmatrix} e^{-3t} & 0 \\ 0 & e^{2t} \end{bmatrix}$. This leads to the solution

$$\mathbf{Y}(t) = \begin{bmatrix} 1 & 11 \\ 1 & 6 \end{bmatrix}\begin{bmatrix} e^{-3t} & 0 \\ 0 & e^{2t} \end{bmatrix} \cdot \frac{1}{5}\begin{bmatrix} -6 & 11 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 11 \\ 1 & 6 \end{bmatrix} \cdot \begin{bmatrix} e^{-3t}(11y_0 - 6x_0) \\ e^{2t}(x_0 - y_0) \end{bmatrix}.$$

$$= \frac{(11y_0 - 6x_0)}{5} \cdot e^{-3t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{(x_0 - y_0)}{5} \cdot e^{2t}\begin{bmatrix} 11 \\ 6 \end{bmatrix}$$

Using the book's initial values $x_0 = 0$ and $y_0 = -5$ we get $\mathbf{Y}(t) = -11e^{-3t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + e^{2t}\begin{bmatrix} 11 \\ 6 \end{bmatrix}$. Thus we have derived $x(t) = 11e^{2t} - 11e^{-3t}$ and $y(t) = 6e^{2t} - 11e^{-3t}$. That agrees with the book's answer. Notice that the prior equation gives us the general solution directly in terms of the initial values x_0 and y_0.

Example 2. On page 298 of the text, we find an example with $A=\begin{bmatrix}-2 & -3\\ 3 & -2\end{bmatrix}$. The eigenvalues are $\lambda=-2+3i$ and $\mu=-2-3i$ with corresponding eigenvectors $\mathbf{v}=\begin{bmatrix}i\\ 1\end{bmatrix}$ and $\mathbf{w}=\begin{bmatrix}-i\\ 1\end{bmatrix}$. We put these into the matrix $P=\begin{bmatrix}i & -i\\ 1 & 1\end{bmatrix}$ and find $P^{-1}=\frac{1}{2i}\begin{bmatrix}1 & i\\ -1 & i\end{bmatrix}$. Also, our diagonal matrix is $D=\begin{bmatrix}-2+3i & 0\\ 0 & -2-3i\end{bmatrix}$, so $e^{Dt}=\begin{bmatrix}e^{(-2+3i)t} & 0\\ 0 & e^{(-2-3i)t}\end{bmatrix}$. This leads to the solution

$$\mathbf{Y}(t)=\begin{bmatrix}i & -i\\ 1 & 1\end{bmatrix}\begin{bmatrix}e^{(-2+3i)t} & 0\\ 0 & e^{(-2-3i)t}\end{bmatrix}\cdot\frac{1}{2i}\begin{bmatrix}1 & i\\ -1 & i\end{bmatrix}\cdot\begin{bmatrix}x_0\\ y_0\end{bmatrix}=\frac{-i}{2}\begin{bmatrix}i & -i\\ 1 & 1\end{bmatrix}\cdot\begin{bmatrix}e^{(-2+3i)t}(x_0+iy_0)\\ e^{(-2-3i)t}(-x_0+iy_0)\end{bmatrix}.$$

$$=\frac{(y_0-ix_0)}{2}\cdot e^{(-2+3i)t}\begin{bmatrix}i\\ 1\end{bmatrix}+\frac{(y_0+ix_0)}{2}\cdot e^{(-2-3i)t}\begin{bmatrix}-i\\ 1\end{bmatrix}$$

This is straightforward, though it does give us a solution in terms of complex numbers. But observe that with real numbers for x_0 and y_0, the two terms above are complex conjugates. Looking just at the first term, we compute

$$\frac{(y_0-ix_0)}{2}\cdot e^{(-2+3i)t}\begin{bmatrix}i\\ 1\end{bmatrix}=\frac{e^{-2t}}{2}\cdot(y_0-ix_0)(\cos(3t)-i\sin(3t))\begin{bmatrix}i\\ 1\end{bmatrix}=\frac{e^{-2t}}{2}\cdot\begin{bmatrix}(x_0+iy_0)(\cos(3t)-i\sin(3t))\\ (y_0-ix_0)(\cos(3t)-i\sin(3t))\end{bmatrix}.$$

Adding this to its complex conjugate we get twice the real part. That leads us to

$$\mathbf{Y}(t)=e^{-2t}\begin{bmatrix}x_0\cos(3t)+y_0\sin(3t)\\ y_0\cos(3t)-x_0\sin(3t)\end{bmatrix}=e^{-2t}\cos(3t)\begin{bmatrix}x_0\\ y_0\end{bmatrix}+e^{-2t}\sin(3t)\begin{bmatrix}y_0\\ -x_0\end{bmatrix}.$$

This is again a general solution to the system expressed directly in terms of the initial values x_0 and y_0. Given a specific initial point, say $(x_0, y_0) = (1,1)$ we can immediately write $\mathbf{Y}(t)=e^{-2t}\cos(3t)\begin{bmatrix}1\\ 1\end{bmatrix}+e^{-2t}\sin(3t)\begin{bmatrix}1\\ -1\end{bmatrix}=e^{-2t}\begin{bmatrix}\cos(3t)+\sin(3t)\\ \cos(3t)-\sin(3t)\end{bmatrix}$, from which we can read off the component functions $x(t) = e^{-2t}(\cos(3t) + \sin(3t))$ and $y(t) = e^{-2t}(\cos(3t) - \sin(3t))$.

Example 3. On page 322 in the text, a system is discussed with $A=\begin{bmatrix}-5 & 1\\ -1 & -3\end{bmatrix}$. The characteristic equation for A is $\lambda^2 + 8\lambda + 16 = (\lambda + 4)^2 = 0$, so there is a single eigenvalue, -4, that is repeated as a root of the equation. Solving the system $A\mathbf{v} = -4\mathbf{v}$, we find that $\mathbf{v}$ can be any multiple of (1,1). So we take $\mathbf{v}$ = (1,1) as our eigenvector for $\lambda = -4$. Now we are supposed to find a matrix P so that $A=PJP^{-1}$ where $J=\begin{bmatrix}-4 & 1\\ 0 & -4\end{bmatrix}$. Multiplying both sides of the equation by P gives us the equivalent equation $AP=PJ$. If the columns of P are **a** and **b**, our desired equation is $A[\mathbf{a}\ \mathbf{b}]=[\mathbf{a}\ \mathbf{b}]\begin{bmatrix}-4 & 1\\ 0 & -4\end{bmatrix}$. Now remember that we can multiply matrices one column at a time. So the previous equation reduces to two single column equations: $A\mathbf{a}=[\mathbf{a}\ \mathbf{b}]\begin{bmatrix}-4\\ 0\end{bmatrix}$ and $A\mathbf{b}=[\mathbf{a}\ \mathbf{b}]\begin{bmatrix}1\\ -4\end{bmatrix}$. On the other hand, matrix-vector multiplication can be interpretted as forming a linear combination of the columns of the matrix, using the entries of the vector as coefficients. Applying this idea to our two equations we obtain $A\mathbf{a} = -4\mathbf{a}$ and $A\mathbf{b} = \mathbf{a} - 4\mathbf{b}$. The first of these is satisfied if we take **a** to be an eigenvector of A corresponding to the eigenvalue -4. Therefore, replace **a** with our vector $\mathbf{v} = (1,1)$. The second equation thus becomes $A\mathbf{b} = \mathbf{v} - 4\mathbf{b}$ or equivalently, $(A + 4I)\mathbf{b} = \mathbf{v}$. This is now a purely numerical system: $\left(\begin{bmatrix} -5 & 1 \\ -1 & -3 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\right)\mathbf{b} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. We can see that the rows are identical, so this system is singular, and there are infinitely many solutions. But we only need one. By inspection, we can take $\mathbf{b} = (0,1)$. Then the matrix P that we need is $[\mathbf{a}\ \mathbf{b}] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

As you can verify, $A\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix}$. Therefore, $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}$ and $e^{At} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} e^{\begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix}t} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}$.

In the outline, it is stated that $e^{\begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix}t} = e^{-4t}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$. The solution to IVP is thus given by

$$\mathbf{Y}(t) = e^{At}\mathbf{Y}_0 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} e^{\begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix}t} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}\mathbf{Y}_0 = e^{-4t}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}\mathbf{Y}_0 .$$

Once again, the matrix approach has led us to a general solution in terms of the initial values. In particular, we can determine a specific solution as soon as we know the initial value.

The general solution above can be related to the general solutions we found in the case of distinct eigenvalues, as follows. For a given specific vector $\mathbf{Y}_0$, the product $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}\mathbf{Y}_0$ will be a particular numerical vector. Let us write that as (k_1, k_2). Then our solution can be rewritten in the form

$$\mathbf{Y}(t) = e^{-4t}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = e^{-4t}\begin{bmatrix} 1 & t \\ 1 & (t+1) \end{bmatrix}\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = k_1 \cdot e^{-4t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + k_2 \cdot e^{-4t}\begin{bmatrix} t \\ t+1 \end{bmatrix}.$$

Here we see a linear combination of two specific solutions, $e^{-4t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $e^{-4t}\begin{bmatrix} t \\ t+1 \end{bmatrix}$. Note that the first is the familiar form we associate with an eigenvalue and its eigenvector. The other is something we have not seen before, and it provides us a second solution when we have only one eigenvalue. The two solutions are linearly independent at time $t = 0$, showing that we can combine them to satisfy any initial condition. In this form, the general solution looks very similar to the ones we found earlier for the case of two distinct eigenvalues.

Exercises

For each system below, express the general solution in the form $P\begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{bmatrix}P^{-1} \cdot \mathbf{Y}_0$ (for distinct eigenvalues) or $e^{\lambda t} P\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}P^{-1} \cdot \mathbf{Y}_0$ (for a repeated eigenvalue), as shown in the examples. Then substitute the specified initial condition and simply the result to something of the form $\mathbf{Y}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$. For complex eigenvalues, this last step may be algebraically too demanding, so don't waste too much time trying to make it work out. You may find this formula useful: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

1. $A = \begin{bmatrix} -5 & -2 \\ -1 & -4 \end{bmatrix}$, (Hint: eigenvalues are -3 and -6, eigenvectors (1,-1) and (2,1).) Initial point is (2,-5). (Compare with problem 3.2.3)

2. $A = \begin{bmatrix} 9 & 16 \\ -9 & 9 \end{bmatrix}$, (Hint: eigenvalues are $9 \pm 12i$, eigenvectors $(4, \pm 3i)$.) Initial point is (4,9). (Solution: $\mathbf{Y}(t) = e^{9t}\cos(12t)\begin{bmatrix} 4 \\ 9 \end{bmatrix} + e^{9t}\sin(12t)\begin{bmatrix} 12 \\ -3 \end{bmatrix}$.)

3. $A = \begin{bmatrix} -2 & -1 \\ 1 & -4 \end{bmatrix}$, (Hint: eigenvalues are -3 and -3, eigenvector (1,1).) Initial point is (1,0). (Compare with problem 3.5.7).

Derivations

From the equation $e^{At} = I + (At) + \dfrac{(At)^2}{2} + \dfrac{(At)^3}{3 \cdot 2} + \dfrac{(At)^4}{4 \cdot 3 \cdot 2} + \cdots$ we can differentiate term by term. That leads to $\left(e^{At}\right)' = 0 + (A) + \dfrac{2A^2 t}{2} + \dfrac{3A^3 t^2}{3 \cdot 2} + \dfrac{4A^4 t^3}{4 \cdot 3 \cdot 2} + \cdots$. Factor out an A and simplify, and you find that $\left(e^{At}\right)' = Ae^{At}$. Now suppose we define a function by $\mathbf{Y}(t) = e^{At}\,\mathbf{Y}_0$. Then $\mathbf{Y}(0) = \mathbf{Y}_0$ and $\mathbf{Y}'(t) = Ae^{At}\,\mathbf{Y}_0 = A\mathbf{Y}(t)$. This verifies that $\mathbf{Y}(t) = e^{At}\,\mathbf{Y}_0$ is the solution to IVP.

In the case of distinct eigenvalues λ and μ, with eigenvectors **v** and **w,** we have $A\mathbf{v} = \lambda\mathbf{v}$ and $A\mathbf{w} = \mu\mathbf{w}$. This can be written in matrix notation as $A\,[\mathbf{v} \quad \mathbf{w}] = [\lambda\mathbf{v} \quad \mu\mathbf{w}]$.

By direct computation we can also verify that $[\mathbf{v}\ \ \mathbf{w}]\begin{bmatrix}\lambda & 0\\ 0 & \mu\end{bmatrix} = [\lambda\mathbf{v}\ \ \mu\mathbf{w}]$. Thus we have $A[\mathbf{v}\ \ \mathbf{w}] = [\mathbf{v}\ \ \mathbf{w}]\begin{bmatrix}\lambda & 0\\ 0 & \mu\end{bmatrix}$. Defining $D = \begin{bmatrix}\lambda & 0\\ 0 & \mu\end{bmatrix}$ and $P = [\mathbf{v}\ \ \mathbf{w}]$, this can be written $AP=PD$. Noting that eigenvectors for distinct eigenvalues are independent, we see that P^{-1} exists, and therefore $A = PDP^{-1}$, as stated in step 5 of the outline.

To compute e^{At} using this result, observe that $At = PDtP^{-1}$. Now look at powers of At. For example, $(At)^3 = (PDtP^{-1})(PDtP^{-1})(PDtP^{-1}) = P(Dt)^3 P^{-1}$. Using this in the power series $e^{At} = I + (At) + \frac{(At)^2}{2} + \frac{(At)^3}{3\cdot 2} + \frac{(At)^4}{4\cdot 3\cdot 2} + \cdots$ shows that $e^{At} = P(e^{Dt})\,P^{-1}$. Moreover, since $Dt = \begin{bmatrix}\lambda t & 0\\ 0 & \mu t\end{bmatrix}$, it is easy to verify that $(Dt)^k = \begin{bmatrix}(\lambda t)^k & 0\\ 0 & (\mu t)^k\end{bmatrix}$. Combining this with the power series $e^{Dt} = I + (Dt) + \frac{(Dt)^2}{2} + \frac{(Dt)^3}{3\cdot 2} + \frac{(Dt)^4}{4\cdot 3\cdot 2} + \cdots$ leads to

$$e^{Dt} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}(\lambda t) & 0\\ 0 & (\mu t)\end{bmatrix} + \frac{1}{2}\begin{bmatrix}(\lambda t)^2 & 0\\ 0 & (\mu t)^2\end{bmatrix} + \frac{1}{3!}\begin{bmatrix}(\lambda t)^3 & 0\\ 0 & (\mu t)^3\end{bmatrix} + \cdots$$

Multiplying the coefficient into each matrix, and then adding them up produces a diagonal matrix with a power series in each diagonal entry. This shows that $e^{Dt} = \begin{bmatrix}e^{\lambda t} & 0\\ 0 & e^{\mu t}\end{bmatrix}$ and so $e^{At} = Pe^{Dt}P^{-1} = P\begin{bmatrix}e^{\lambda t} & 0\\ 0 & e^{\mu t}\end{bmatrix}P^{-1}$. This justifies the final conclusion of step 5 of the outline.

Proceeding to step 6, we must now consider the case of a matrix with a single eigenvalue. In this case, we cannot diagonalize the matrix, but the outline says we can triangularize it. This was actually done in detail in Example 3. The method shown there applies in general. Say the unique eigenvalue is λ, with eigenvector **v**. So we know that $A\mathbf{v} = \lambda\mathbf{v}$. The goal is to find an independent vector **b** such that $A[\mathbf{v}\ \ \mathbf{b}] = [\mathbf{v}\ \ \mathbf{b}]\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$. As shown in the example, we can find **b** by solving $A\mathbf{b} = \mathbf{v} + \lambda\mathbf{b}$ which in turn reduces to $(A - \lambda I)\mathbf{b} = \mathbf{v}$. Such an equation is always solvable when there is a unique eigenvalue (except for the trivial case where $A = \lambda I$). This is stated in the book (top of page 319) and there is an outline of the proof in problem 3.5.16. The point of all of this is that it is always possible to triangularize A in the case of a single eigenvalue, and the steps above give a method to do so.

That leaves us to compute the exponential matrix $e^{\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}t}$. Following are the detailed steps for Example 3. The general case is essentially the same.

To find $e^{\begin{bmatrix}-4 & 1\\ 0 & -4\end{bmatrix}t}$, first write $\begin{bmatrix}-4 & 1\\ 0 & -4\end{bmatrix}t$ as $\begin{bmatrix}-4t & 0\\ 0 & -4t\end{bmatrix}+\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}$. By the laws of exponents, $e^{\begin{bmatrix}-4t & 0\\ 0 & -4t\end{bmatrix}+\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}} = e^{\begin{bmatrix}-4t & 0\\ 0 & -4t\end{bmatrix}}e^{\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}}$. This can be verified using the series definition of the matrix exponential, and the fact that $\begin{bmatrix}-4t & 0\\ 0 & -4t\end{bmatrix}\cdot\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}=\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}\cdot\begin{bmatrix}-4t & 0\\ 0 & -4t\end{bmatrix}$. Continuing, we know $e^{\begin{bmatrix}-4t & 0\\ 0 & -4t\end{bmatrix}}=\begin{bmatrix}e^{-4t} & 0\\ 0 & e^{-4t}\end{bmatrix}$ so we need to find $e^{\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}}$. Here we use the series definition directly: $e^{\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}+\frac{1}{2}\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}^2+\frac{1}{3\cdot 2}\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}^3+\cdots$. Matrix multiplication shows that $\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}^2=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$ and so the same is true for all the higher powers. Consequently, $e^{\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\begin{bmatrix}0 & t\\ 0 & 0\end{bmatrix}=\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$ and we can now compute $e^{\begin{bmatrix}-4 & 1\\ 0 & -4\end{bmatrix}t}=\begin{bmatrix}e^{-4t} & 0\\ 0 & e^{-4t}\end{bmatrix}\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}=e^{-4t}\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$.

By an exactly analogous process, it can be shown that, for any λ, $e^{\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}t}=e^{\lambda t}\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$. This completes the derivation of step 6 of the outline.