Solving 1st-order linear vector-valued ODE's

Let's look for solutions to the unforced system:

$$\dot{X} = AX \tag{1}$$

Suppose that V, λ is an eigenvector-eigenvalue pair of A. In other words, multiplying V by matrix A is the same as multiplying V by the eigenvalue, λ:

$$AV = \lambda V \tag{2}$$

As a consequence of this, $X(t) = Ve^{\lambda t}$ is a valid solution to the unforced system . To prove this, let's evaluate both $\dot{X}$ and AX, and then compare their results. Beginning with $\dot{X}$, we get:

$$\dot{X} = \frac{d}{dt}\left(Ve^{\lambda t}\right) = \lambda V e^{\lambda t} \tag{3}$$

Evaluating the product AX:

$$AX = A(Ve^{\lambda t}) = (AV)e^{\lambda t} = \lambda V e^{\lambda t} \tag{4}$$

From this, we see that $\dot{X}$ and AX are equal to one another, meaning that $X(t)$ is a solution to the differential equation:

$$\boxed{\dot{X} = \lambda V e^{\lambda t} = AX} \tag{5}$$

Now that we have identified a single type of solution ($X(t) = Ve^{\lambda t}$ where V and λ are an eigenvector-eigenvalue pair), let's figure out a way to generalize this result to construct a larger space of solutions!

Proof of linearity: Suppose that we have identified two different solutions, $X_1(t)$ and $X_2(t)$ to our differential equation. In other words, $X_1(t)$ and $X_2(t)$ both satisfy:

$$\dot{X}_1 = AX_1, \quad \dot{X}_2 = AX_2 \tag{6}$$

Let's define function $X_3(t)$ as a linear combination of $X_1(t)$ and $X_2(t)$:

$$X_3(t) = \alpha_1 X_1(t) + \alpha_2 X_2(t) \tag{7}$$

It turns out that $X_3(t)$ is also a solution to the differential equation. To see this, we can take a linear combination of the equations that X_1 and X_2 satisfy:

$$\alpha_1\left(\dot{X}_1 = AX_1\right) + \alpha_2\left(\dot{X}_2 = AX_2\right) \quad \rightarrow \quad \alpha_1\dot{X}_1 + \alpha_2\dot{X}_2 = \alpha_1 AX_1 + \alpha_2 AX_2 \tag{8}$$

Rearranging the right side of the equation, we get:

$$\alpha_1\dot{X}_1 + \alpha_2\dot{X}_2 = A\left(\alpha_1 X_1 + \alpha_2 X_2\right) \quad \rightarrow \quad \dot{X}_3 = AX_3 \tag{9}$$

which means that X_3 is indeed a solution of the differential equation. This property of the ODE (being able to combine solutions to generate new ones) is called **superposition**. Let's now construct a general form for solutions to 2.

Suppose that we have identified n eigenvalue-eigenvector pairs of A, (V_i, λ_i):

$$AV_1 = \lambda_1 V_1, \quad AV_2 = \lambda_2 V_2, \quad ..., \quad AV_n = \lambda_n V_n, \tag{10}$$

Let's define $X(t)$ as a linear combination of the building block solutions that we identified ($V_i e^{\lambda_i t}$):

$$X(t) = \sum_{i=1}^{n} \alpha_i V_i e^{\lambda_i t} = \alpha_1 V_1 e^{\lambda_1 t} + \alpha_2 V_2 e^{\lambda_2 t} + ... + \alpha_n V_n e^{\lambda_n t} \tag{11}$$

From the superposition property that we just proved, $X(t)$ must be a solution to 2. Alternatively, we can just do the math and compare $\dot{X}$ with AX. Evaluating $\dot{X}$, we get:

$$\dot{X}(t) = \frac{d}{dt}\left(\sum_{i=1}^{n} \alpha_i V_i e^{\lambda_i t}\right) = \frac{d}{dt}\left(\alpha_1 V_1 e^{\lambda_1 t} + \alpha_2 V_2 e^{\lambda_2 t} + ... + \alpha_n V_n e^{\lambda_n t}\right) \tag{12}$$

$$\dot{X}(t) = \sum_{i=1}^{n} \alpha_i \lambda_i V_i e^{\lambda_i t} = \alpha_1 \lambda_1 V_1 e^{\lambda_1 t} + \alpha_2 \lambda_2 V_2 e^{\lambda_2 t} + ... + \alpha_n \lambda_n V_n e^{\lambda_n t} \tag{13}$$

Evaluating AX, we get:

$$AX(t) = A\left(\sum_{i=1}^{n} \alpha_i V_i e^{\lambda_i t}\right) = A\left(\alpha_1 V_1 e^{\lambda_1 t} + \alpha_2 V_2 e^{\lambda_2 t} + ... + \alpha_n V_n e^{\lambda_n t}\right) \tag{14}$$

Since matrix multiplication distributes across addition and commutes with scalar multiplication, we see that:

$$AX(t) = \sum_{i=1}^{n} \alpha_i (AV_i) e^{\lambda_i t} = \alpha_1 (AV_1) e^{\lambda_1 t} + \alpha_2 (AV_2) e^{\lambda_2 t} + ... + \alpha_n (AV_n) e^{\lambda_n t} \tag{15}$$

Substituting in $\lambda_i V_i = AV_i$ gives us:

$$AX(t) = \sum_{i=1}^{n} \alpha_i \lambda_i V_i e^{\lambda_i t} = \alpha_1 \lambda_1 V_1 e^{\lambda_1 t} + \alpha_2 \lambda_2 V_2 e^{\lambda_2 t} + ... + \alpha_n \lambda_n V_n e^{\lambda_n t} \tag{16}$$

From this, we see that $\dot{X}$ and AX are equivalent:

$$\dot{X}(t) = \sum_{i=1}^{n} \alpha_i \lambda_i V_i e^{\lambda_i t} = AX(t) \tag{17}$$

meaning that $X(t)$ is a solution to 2. Our expression for $X(t)$ is now general enough for us to be able to solve the corresponding initial value problem (IVP):

$$\dot{X} = AX, \qquad X(t=0) = X_0 \tag{18}$$

Let's assume that X is an n-dimensional vector, and that the n eigenvectors that we have identified and chosen, $(V_1, V_2, ..., V_n)$ are linearly independent (in other words, A is diagonalizable). In this case, let's see if we can find a way to get the general form of $X(t)$ to satisfy the initial conditions. To do this, let's first evaluate $X(t)$ at $t = 0$:

$$X(t=0) = \sum_{i=1}^{n} \alpha_i V_i e^{\lambda_i 0} = \alpha_1 V_1 e^{\lambda_1 0} + \alpha_2 V_2 e^{\lambda_2 0} + ... + \alpha_n V_n e^{\lambda_n 0} \tag{19}$$

Since $e^0 = 1$, we see that $X(t=0)$ reduces to:

$$X(t=0) = X_0 = \sum_{i=1}^{n} \alpha_i V_i = \alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_n V_n \tag{20}$$

We can express this linear combination as a matrix-vector product:

$$X(t=0) = X_0 = \sum_{i=1}^{n} \alpha_i V_i = \alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_n V_n \tag{21}$$

$$X_0 = \underbrace{\left[\ V_1 \mid V_2 \mid ... \mid V_n\ \right]}_{M} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = M \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{22}$$

Here, the vertical bar, $|$, indicates horizontal concatenation. In other words, the ith column of matrix M is the ith eigenvector, V_i. Since we assumed that the eigenvectors we chose were linearly independent, M is an invertible n x n matrix. This means that we can solve for the vector of coefficients, $[\alpha_1, \alpha_2, ..., \alpha_n]^T$ that allow us to satisfy the initial conditions:

$$M^{-1} X_0 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{23}$$

In summary, we can use the following process to solve the IVP described by 18:

1. Assuming that A is diagonalizable identify n linearly independent eigenvectors, $(V_1, V_2, ..., V_n)$ and their corresponding eigenvalues $(\lambda_1, \lambda_2, ..., \lambda_n)$. You can compute this numerically in MATLAB using the eig function:

2. Construct the matrix of eigenvectors, $M = [V_1|V_2|...|V_n]$. (In MATLAB, this is automatically one of the outputs of eig.)

3. Find the vector of coefficients $[\alpha_1, \alpha_2, ..., \alpha_n]^T$ that allows the solution to satisfy the initial conditions:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = M^{-1} X_0 \tag{24}$$

4. Now that we have found the eigenvectors, eigenvalues, and summation coefficients, the solution to the IVP, $X(t)$, is given by:

$$X(t) = \sum_{i=1}^{n} \alpha_i V_i e^{\lambda_i t} = \alpha_1 V_1 e^{\lambda_1 t} + \alpha_2 V_2 e^{\lambda_2 t} + ... + \alpha_n V_n e^{\lambda_n t} \tag{25}$$

Computing the eigenvectors and eigenvalues of a matrix

At this point, it is a worthwhile exercise to review the process of how to compute the eigenvectors and eigenvalues of matrix A. Suppose that V and λ are an eigenvector-eigenvalue pair of n x n matrix A:

$$AV = \lambda V \tag{26}$$

We can make the substitution $V = IV$ where I is the n x n identity matrix:

$$AV = \lambda IV \tag{27}$$

Moving terms to the same side, we get:

$$AV - \lambda IV = 0 \tag{28}$$

Since matrix-vector multiplication distributes across addition, we can factor out the V from the left side of the equation, giving us:

$$(A - \lambda I)\, V = 0 \tag{29}$$

In other words, V belongs to the nullspace of the matrix $(A - \lambda I)$. For $(A - \lambda I)$ to have a nontrivial nullspace (a nullspace that isn't just the vector of zeros), its determinant must equal zero:

$$\det (A - \lambda I) = 0 \tag{30}$$

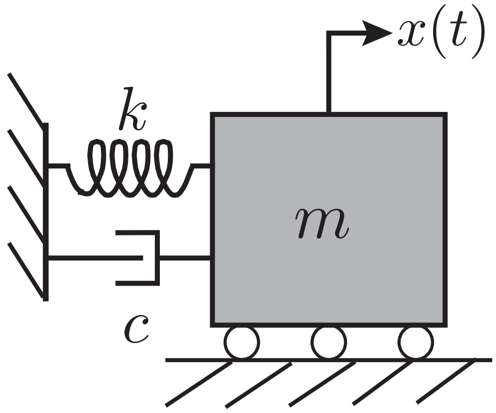
The left-hand side of the equation is an n-degree polynomial in λ called the characteristic polynomial. We can then solve for the values of λ by finding the roots of this polynomial. To find the corresponding eigenvector, V, we must substitute λ back into the previous equation:

$$(A - \lambda I)\, V = 0 \tag{31}$$

The eigenvector, V, can then be computed as a nullspace vector of $(A - \lambda I)$. To see how this all works, let's look at an example:

Example: Mass-Spring-Damper System Revisited

Let's examine our old friend, the unforced mass-spring-damper system:



The 2nd-order differential equation describing the motion of this system is given by:

$$m\ddot{x} + c\dot{x} + kx = 0 \tag{32}$$

We can rewrite this 2nd-order ODE as a 1st-order vector-valued ODE by introducing the variable $v = \dot{x}$ which is the velocity of the cart. The vector-valued state, X is the combination of the position, x, and velocity, v:

$$X(t) = \begin{bmatrix} x(t) \\ v(t) \end{bmatrix} \tag{33}$$

Taking the derivative of the state vector, $\dot{X}$, we get:

$$\dot{X} = \begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \ddot{x} \end{bmatrix} \tag{34}$$

Note that we can arrange 32 to express the acceleration, $\ddot{x}$ in terms of x and $\dot{x}$:

$$\ddot{x} = -\frac{k}{m}x - \frac{c}{m}\dot{x} \tag{35}$$

Substituting v and $\dot{v}$ for $\dot{x}$ and $\ddot{x}$ respectively, we get:

$$\dot{v} = -\frac{k}{m}x - \frac{c}{m}v \tag{36}$$

This gives us two 1st-order differential equations describing the motion of the system:

$$\dot{x} = v \tag{37}$$

$$\dot{v} = -\frac{k}{m}x - \frac{c}{m}v \tag{38}$$

which we can express using matrix-vector multiplication:

$$\underbrace{\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix}}_{\dot{X}} = \underbrace{\begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ v \end{bmatrix}}_{X} \quad \rightarrow \quad \dot{X} = AX \tag{39}$$

To find solutions to this ODE, we must first identify eigenvector-eigenvalue pairs (V_i, λ_i). In the previous section, we showed that the eigenvalues must satisfy the equation:

$$\det(A - \lambda I) = 0 \tag{40}$$

Plugging in for A into the left side of the equation, we get:

$$0 = \det(A - \lambda I) = \det\left(\begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} -\lambda, & 1 \\ -\frac{k}{m}, & -\frac{c}{m} - \lambda \end{bmatrix}\right) \tag{41}$$

We can now evaluate the determinant to generate the characteristic polynomial:

$$(-\lambda)\left(-\frac{c}{m}-\lambda\right)-\left(-\frac{k}{m}\right)(1)=0 \tag{42}$$

Simplifying, we get:

$$\lambda^2+\frac{c}{m}\lambda+\frac{k}{m}=0 \tag{43}$$

or equivalently:

$$\lambda^2+2\zeta\omega_n\lambda+\omega_n^2=0, \qquad \omega_n=\sqrt{\frac{k}{m}}, \quad \zeta=\frac{c}{2\sqrt{km}} \tag{44}$$

Note that this polynomial is identical to the one we have always generated when searching for solutions to the 2nd-order ODE. Thus, our eigenvalues are given by:

$$\lambda=-\zeta\omega_n\pm\omega_n\sqrt{\zeta^2-1} \tag{45}$$

which may be real or complex, depending on whether or not the system is over/under-damped.

Matrix Exponentiation

You may recall that if A is a diagonalizable matrix, then it can be written in the form:

$$A=MDM^{-1} \tag{46}$$

where M is the matrix whose columns correspond to the eigenvectors of A:

$$M=[V_1,V_2,...,V_i,...,V_n] \tag{47}$$

and D is a diagonal matrix whose diagonal elements are the corresponding eigenvalues of A:

$$D=\begin{bmatrix}\lambda_1 & 0 & ... & 0\\ 0 & \lambda_2 & ... & 0\\ \vdots & \vdots & \lambda_i & \vdots\\ 0 & 0 & ... & \lambda_n\end{bmatrix} \tag{48}$$

This matrix factorization is called the eigendecomposition of A. We can use the eigendecomposition to express solutions to the IVP described by 18 more succinctly:

$$\dot{X}=AX, \qquad X(t=0)=X_0 \tag{49}$$

We just derived that the solution to the IVP, $X(t)$, can be written in the form:

$$X(t)=\sum_{i=1}^{n}\alpha_iV_ie^{\lambda_it}=\alpha_1V_1e^{\lambda_1t}+\alpha_2V_2e^{\lambda_2t}+...+\alpha_nV_ne^{\lambda_nt} \tag{50}$$

where the coefficients α_i can be found by computing:

$$\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}=M^{-1}X_0 \tag{51}$$

We can rewrite the expression for $X(t)$ in equation 50 as matrix-vector multiplication:

$$X(t)=\sum_{i=1}^{n}\alpha_iV_ie^{\lambda_it}=\underbrace{\left[\begin{array}{c|c|c|c} V_1 & V_2 & ... & V_n\end{array}\right]}_{M}\begin{bmatrix}\alpha_1e^{\lambda_1t}\\ \alpha_2e^{\lambda_2t}\\ \vdots\\ \alpha_ne^{\lambda_nt}\end{bmatrix} \tag{52}$$

Note that the matrix in this product is the matrix of eigenvectors, M. Furthermore, we can factor the vector in this product as follows:

$$\begin{bmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_n e^{\lambda_n t} \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & e^{\lambda_i t} & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{53}$$

Substituting this into our expression for $X(t)$, we get:

$$X(t) = M \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & e^{\lambda_i t} & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{54}$$

Substituting in $M^{-1}X_0$ for $[\alpha_1, .., \alpha_n]^T$, we get:

$$X(t) = \underbrace{\left(M \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & e^{\lambda_i t} & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix} M^{-1} \right)}_{e^{At}} X_0 \tag{55}$$

$$\boxed{X(t) = e^{At} X_0} \tag{56}$$

The matrix, e^{At} is the matrix exponential of At. There are several ways to compute the matrix exponential. One such method is as follows:

1. Diagonalize A:
$$A = MDM^{-1} \tag{57}$$

2. compute a new diagonal matrix, $Q = e^{Dt}$, where ith diagonal element of Q is e to the power of $\lambda_i t$ where λ_i is the ith eigenvalue of A:
$$Q = e^{Dt} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & e^{\lambda_i t} & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix} \tag{58}$$

3. Multiply M, Q, and M^{-1} to get the desired matrix exponential:
$$e^{At} = MQM^{-1} = Me^{Dt}M^{-1} \tag{59}$$

An equivalent representation of e^{At} is to use the power series expansion of e^x:

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \dots \tag{60}$$

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} (At)^k = 1 + (At) + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \frac{1}{24}(At)^4 + \dots \tag{61}$$

How and why are these two forms equivalent? To see this, we will use the eigendecomposition of A, $A = MDM^{-1}$. We can prove by induction that $(At)^n$ can be factored into the following product:

$$(At)^n = M\,(Dt)^n\,M^{-1} \tag{62}$$

This is true for the base cases of $n = 0$ and $n = 1$, since:

$$I = (At)^0 = M\,(Dt)^0\,M^{-1} = M\,(I)\,M^{-1} = MM^{-1} = I \tag{63}$$

$$At = (At)^1 = M\,(Dt)^1\,M^{-1} = M(Dt)M^{-1} = At \tag{64}$$

To perform the inductive step, assume that this formula is true for $n = k$. It follows that:

$$(At)^{k+1} = (At)^k(At) = M(Dt)^kM^{-1}M(Dt)M^{-1} = M(Dt)^k(Dt)M^{-1} = M(Dt)^{k+1}M^{-1} \tag{65}$$

This means that if the formula is true for $n = k$, then it is also true for $n = k + 1$ which completes the proof. We can now plug this formula into the power series representation of e^{At}:

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!}(At)^k = \sum_{k=0}^{\infty} \frac{1}{k!}M(Dt)^kM^{-1} \tag{66}$$

Since matrix multiplication distributes across addition, M and M^{-1} can be factored out of the summation:

$$e^{At} = M\left(\sum_{k=0}^{\infty}\frac{1}{k!}(Dt)^k\right)M^{-1} = M\left(1 + (Dt) + \frac{1}{2}(Dt)^2 + \frac{1}{6}(Dt)^3 + \frac{1}{24}(Dt)^4 + ...\right)M^{-1} \tag{67}$$

Since D is the diagonal matrix of the eigenvalues, we can rewrite the summation in matrix form:

$$Dt = \begin{bmatrix} \lambda_1 t & 0 & ... & 0 \\ 0 & \lambda_2 t & ... & 0 \\ \vdots & \vdots & \lambda_i t & \vdots \\ 0 & 0 & ... & \lambda_n t \end{bmatrix} \tag{68}$$

$$\rightarrow \sum_{k=0}^{\infty}\frac{1}{k!}(Dt)^k = \begin{bmatrix} \sum_{k=0}^{\infty}\frac{1}{k!}(\lambda_1 t)^k & 0 & ... & 0 \\ 0 & \sum_{k=0}^{\infty}\frac{1}{k!}(\lambda_2 t)^k & ... & 0 \\ \vdots & \vdots & \sum_{k=0}^{\infty}\frac{1}{k!}(\lambda_i t)^k & \vdots \\ 0 & 0 & ... & \sum_{k=0}^{\infty}\frac{1}{k!}(\lambda_n t)^k \end{bmatrix} \tag{69}$$

Note that the ith diagonal element is the power series expansion of $e^{\lambda_i t}$:

$$\sum_{k=0}^{\infty}\frac{1}{k!}(\lambda_i t)^k = e^{\lambda_i t} \tag{70}$$

Substituting this into our expression for the summation, we get:

$$\sum_{k=0}^{\infty}\frac{1}{k!}(Dt)^k = \begin{bmatrix} e^{\lambda_1 t} & 0 & ... & 0 \\ 0 & e^{\lambda_2 t} & ... & 0 \\ \vdots & \vdots & e^{\lambda_i t} & \vdots \\ 0 & 0 & ... & e^{\lambda_n t} \end{bmatrix} \tag{71}$$

Substituting this into our power series expansion for e^{At}, we get:

$$e^{At} = M\left(\sum_{k=0}^{\infty}\frac{1}{k!}(Dt)^k\right)M^{-1} = M\begin{bmatrix} e^{\lambda_1 t} & 0 & ... & 0 \\ 0 & e^{\lambda_2 t} & ... & 0 \\ \vdots & \vdots & e^{\lambda_i t} & \vdots \\ 0 & 0 & ... & e^{\lambda_n t} \end{bmatrix}M^{-1} \tag{72}$$

which was our original expression for e^{At}. It should be noted that MATLAB has its own built-in matrix exponentiation function, expm, that you can use.