Solutions to Challenge Problem Set 1, Math 292 Spring 2018

Eric A. Carlen[1]
Rutgers University

February 26, 2018

This challenge problem set concerns the simplest system of second order differential equations

$$\begin{aligned} x''(t) &= -x(t) \\ y''(t) &= -y(t) \ . \end{aligned} \tag{0.1}$$

The system is already in decoupled form, and already in Example 12 of Chapter 1 we have solved the equation $x'' + \omega^2 x = 0$. For $\omega = 1$, the case at hand, we have found that the general solution is

$$x(t) = r\sin(t + \theta_0) \ ,$$

for arbitrary $r \geq 0$ and $\theta \in [0, 2\pi)$. Let us seek the solution with

$$x(0) = x_0 \qquad \text{and} \qquad x'(0) = u_0 \ . \tag{0.2}$$

From the general solution we find

$$r\sin(\theta_0) = x_0 \qquad \text{and} \qquad r\cos(\theta_0) = u_0 \ . \tag{0.3}$$

and then from the trigonometric angle addition formulas, we find that the solution of $x'' = -x$ with $x(0) = x_0$ and $x'(0) = u_0$ is

$$x(t) = x_0 \cos t + u_0 \sin t \ .$$

Likewise, the solution of $y'' = -y$ with $y(0) = y_0$ and $y'(0) = v_0$ is

$$y(t) = y_0 \cos t + v_0 \sin t \ .$$

We can combine these solutions as follows: Introduce the vectors $\mathbf{x}(t) = (x(t), y(t))$, $\mathbf{x}_0 = (x_0, y_0)$ and $\mathbf{v}_0 = (u_0, v_0)$.

Then introduce the matrix

$$A := \begin{bmatrix} x_0 & u_0 \\ y_0 & v_0 \end{bmatrix} = [\mathbf{x}_0, \mathbf{v}_0] \ .$$

Then we can summarize our conclusions so far by saying that the unique solution of

$$\mathbf{x}''(t) = -\mathbf{x}(t) \qquad \text{with} \quad \mathbf{x}(0) = \mathbf{x}_0 \quad \text{and} \quad \mathbf{x}'(0) = \mathbf{v}_0 \tag{0.4}$$

is

$$\mathbf{x}(t) = A\mathbf{u}(t) \qquad \text{where} \quad \mathbf{u}(t) = (\cos t, \sin t)\ . \tag{0.5}$$

It may seem that this problem is completely solved, and there is nothing more for the Challenge Workshop to deal with. However, finding the general solution is often only the beginning of the investigation of the solutions of a differential equation. A central focus of the theory is on understanding the *behavior* of the solutions of a differential equation, and even for this simplest of second order systems, much interesting mathematics is connected with understanding the behavior of the solutions.

For example, we can ask: What kind of curves are the parameterized curves given by (0.5)? It turns out that they are always ellipses, though possibly degenerate. This is a consequence of an important general fact: Let B be any invertible 2×2 matrix. Then with $\mathbf{u}(t)$ being the parameterization of the unit circle given in (0.5), $B\mathbf{u}(t)$ is a parameterized ellipse. That is, *the image of the unit circle under any linear transformation from* $\mathbb{R}^2$ *to* $\mathbb{R}^2$ *is an ellipse.*

Exercise 1: Suppose that the matrix A in (0.5) is invertible. Show that with $\mathbf{x}(t)$ given by (0.5),

$$\|A^{-1}\mathbf{x}(t)\|^2 = 1 \tag{0.6}$$

for all t. Define the matrix M by $M = (A^{-1})^T(A^{-1})$, where $(A^{-1})^T$ denote the transpose of A^{-1}. Show also that the equation (0.6) can be written as

$$\mathbf{x}(t) \cdot M\mathbf{x}(t) = 1 \tag{0.7}$$

for all t, where M is a symmetric matrix, and show that if $M = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$, then the equation (0.7) can be written as

$$ax^2(t) + by^2(t) + 2cx(t)y(t) = 1 \tag{0.8}$$

for all t.

SOLUTION Since A is invertible, (0.5) is equivalent to

$$\mathbf{x}(t) = A^{-1}\mathbf{u}(t)\ ,$$

and since $\mathbf{u}(t)$ is a unit vector, then $\|A^{-1}\mathbf{x}(t)\| = 1$ for all t. Thus,

$$1 = \|\|A^{-1}\mathbf{x}(t)\|^2 = (A^{-1}\mathbf{x}(t)) \cdot (A^{-1}\mathbf{x}(t)) = \mathbf{x}(t)) \cdot ((A^{-1})^T A^{-1}\mathbf{x}(t)) = \mathbf{x}(t) \cdot M\mathbf{x}(t)\ .$$

Now simply multiplying out $(x(t0, y(t)) \cdot \begin{bmatrix} a & c \\ c & b \end{bmatrix} (x(t0, y(t))$ we find

$$ax^2(t) + by^2(t) + 2cx(t)y(t) = 1\ .$$

The equation

$$ax^2 + by^2 + 2cxy = 1$$

is the equation of a conic section, and the only bounded conic sections are ellipses. Hence the curve $\mathbf{x}(t)$ must trace out an ellipse, and since there are no linear terms, it is an ellipse centered at the origin.

This raises the question: Which ellipse centered at the origin is it? What are the major and minor axes? These questions are relevant to understanding the behavior of our solution since if R_1 is half the length of the minor axis, and R_2 is half of the length of the major axis,

$$\min_{t\in\mathbb{R}} \|\mathbf{x}(t)\| = R_1 \qquad \text{and} \qquad \max_{t\in\mathbb{R}} \|\mathbf{x}(t)\| = R_2 \ .$$

We know how to find the major and minor axes of an ellipse by writing the equation $ax^2 + by^2 + 2cxy = 1$ in matrix form as

$$(x, y) \cdot \begin{bmatrix} a & c \\ c & b \end{bmatrix} (x, y) = 1 \ ,$$

and then finding the eigenvalues and eigenvectors of M.

Exercise 2: Since M is a symmetric matrix, there exists an orthonormal basis of $\mathbb{R}^2$, $\{\mathbf{u}_1, \mathbf{u}_2\}$ consisting of eigenvectors of M so that for eigenvalues λ_1 and λ_2 we have

$$M\mathbf{u}_1 = \lambda_1\mathbf{u}_1 \quad \text{and} \quad M\mathbf{u}_2 = \lambda_1\mathbf{u}_2 \ .$$

Show that both λ_1 and λ_2 are positive. Changing the labels if need be, we may suppose without loss of generality that

$$\lambda_1 \geq \lambda_2 > 0 \ .$$

Show that the major axis of the ellipse has

$$\pm\frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2$$

as its endpoints, and that the minor axis of the ellipse has

$$\pm\frac{1}{\sqrt{\lambda_1}}\mathbf{u}_1$$

as its endpoints.

Show that if $\mathbf{x}(t)$ is the solution of our system given in (0.5), then $\|\mathbf{x}(t)\|$ is maximal if and only if

$$\mathbf{x}(t) = \pm\frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2 \ ,$$

and prove the corresponding statement about minima.

SOLUTION Since by definition $M = (A^{-1})^T(A^{-1})$, for all non-zero $\mathbf{x} \in \mathbb{R}^2$,

$$\mathbf{x} \cdot M\mathbf{x} = \mathbf{x} \cdot (A^{-1})^T(A^{-1})\mathbf{x} = (A^{-1}\mathbf{x}) \cdot (A^{-1}\mathbf{x}) = \|A^{-1}\mathbf{x}\|^2 > 0 \ .$$

Since we know that all of the eigenvalues of a symmetric matrix B are positive if and only if $\mathbf{x} \cdot B\mathbf{x} > 0$ for all non-zero $\mathbf{x}$, it follows that the eigenvalues of M are both positive. Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be an orthonormal basis of eigenvectors of M with

$$M\mathbf{u}_1 = \lambda_1\mathbf{u}_1 \quad \text{and} \quad M\mathbf{u}_2 = \lambda_1\mathbf{u}_2$$

where $\lambda_1 \geq \lambda_2 > 0$.

Define a system of coordinates on the plane by $u(\mathbf{x}) = \mathbf{u}_1 \cdot \mathbf{x}$ and $v(\mathbf{x}) = \mathbf{u}_2 \cdot \mathbf{x}$ so that $\mathbf{x} = u(\mathbf{x})\mathbf{u}_1 + v(\mathbf{x})\mathbf{u}_2$, and then

$$\begin{aligned} \mathbf{x} \cdot M\mathbf{x} = (u(\mathbf{x})\mathbf{u}_1 + v(\mathbf{x})\mathbf{u}_2) \cdot M(u(\mathbf{x})\mathbf{u}_1 + v(\mathbf{x})\mathbf{u}_2) &= (u(\mathbf{x})\mathbf{u}_1 + v(\mathbf{x})\mathbf{u}_2) \cdot (u(\mathbf{x})\lambda_1\mathbf{u}_1 + v(\mathbf{x})\lambda_2\mathbf{u}_2) \\ &= \lambda_1 u^2(\mathbf{x}) + \lambda_2 v^2(\mathbf{x}) \ . \end{aligned}$$

Thus in terms of the (u, v) cordinate, the points that satisfy $\mathbf{x} \cdot M\mathbf{x}$ have coordinates satisfying

$$\lambda_1 u^2 + \lambda_2 v^2 = 1 \ .$$

This is the euqation of an ellipse whose major axis runs along the v coordonate axis, and whose minor axis runs a long the u cooridnate axis. Since the u axis runs in the u_1 direction, and the v axis runs in the $\mathbf{u}_2$ direction, the endpoints of the major axis are $\pm\frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2$ and the enpoints of the minor axis are $\pm\frac{1}{\sqrt{\lambda_1}}\mathbf{u}_1$.

Sicce we have seen that $\mathbf{x}(t)$ must be on this ellipse, $\|\mathbf{x}(t)\|$ is maximal if and only if $\mathbf{x}(t)$ is an enpoint of the major axis, which means

$$\mathbf{x}(t) = \pm\frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2 \ .$$

Likewise, $\|\mathbf{x}(t)\|$ is minimal if and only if $\mathbf{x}(t)$ is an enpoint of the minor axis, which means

$$\mathbf{x}(t) = \pm\frac{1}{\sqrt{\lambda_1}}\mathbf{u}_1 \ .$$

We have now found the points on the path traced out by the solution $\mathbf{x}(t)$ at which $\|\mathbf{x}(t)\|$ takes on its maximal and minimal values. We now ask: What are the times t at which $\|\mathbf{x}(t)\|$ takes on its maximal and minimal values?

Since $\mathbf{x}(t) = A(\cos t, \sin t)$, to answer this question for the maximal values we have to solve

$$A(\cos t, \sin t) = \pm\frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2$$

and to answer it for minima, we have to solve

$$A(\cos t, \sin t) = \pm\frac{1}{\sqrt{\lambda_1}}\mathbf{u}_1.$$

We know there are solutions. If we multiply both sides of these equations by A^{-1}, the left hand side is a unit vector, and so the right hand side must be a unit vector too. In fact, the two unit vectors we get from these two equations are mutually orthogonal, as the next exercise shows.

Exercise 3. With $\{\mathbf{u}_1, \mathbf{u}_2\}$ and $\{\lambda_1, \lambda_2\}$ defined as in the previous exercise, define

$$\mathbf{v}_1 = \frac{1}{\sqrt{\lambda_1}}A^{-1}\mathbf{u}_1 \quad \text{and} \quad \mathbf{v}_2 = \frac{1}{\sqrt{\lambda_2}}A^{-1}\mathbf{u}_2 \ . \tag{0.9}$$

Show that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for $\mathbb{R}^2$.

Then show that for our solution $\mathbf{x}(t)$, $\|\mathbf{x}(t)\|$ is maximal if and only if t satisfies

$$(\cos t, \sin t) = \pm \mathbf{v}_2 \ ,$$

and prove the corresponding statement for minima.

SOLUTION We check orthonormality by computing 3 dot products: For $j = 1, 2$

$$\|\mathbf{v}_j\|^2 = \mathbf{v}_j \cdot \mathbf{v}_j = \frac{1}{\lambda_j}\mathbf{u}_j \cdot (A^{-1})^T A^{-1}\mathbf{u}_j = \frac{1}{\lambda_j}\mathbf{u}_j \cdot M\mathbf{u}_j = \mathbf{u}_j \cdot \mathbf{u}_j = 1$$

where we have used the fact that $M\mathbf{u}_j = \lambda_j \mathbf{u}_j$. Thus both $\mathbf{v}_1$ and $\mathbf{v}_2$ are unit vectors.

Next, in the same way,

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \frac{1}{\sqrt{\lambda_1\lambda_2}}\mathbf{u}_1 \cdot M\mathbf{u}_2 = \frac{1}{\sqrt{\lambda_1\lambda_2}}\mathbf{u}_1 \cdot \lambda_2\mathbf{u}_2 = 0$$

since $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$. This proves that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is orthonormal.

Now we have seen that $\|\mathbf{x}(t)\|$ is maximal if and only if $\mathbf{x}(t) = \pm\frac{1}{\sqrt{\lambda_2}}\mathbf{u}_2$. Since $\mathbf{x}(t) = A\mathbf{u}(t)$, where $\mathbf{u}(t) = \cos t, \sin t)$, this is the same as

$$\mathbf{u}(t) = \pm\frac{1}{\sqrt{\lambda_2}}A^{-1}\mathbf{u}_2 = \pm\mathbf{v}_2 \ .$$

The same reasoning shows that $\|\mathbf{x}(t)\|$ is minimial if and only if $\mathbf{u}(t) = \pm\mathbf{v}_1$.

Now let us apply what we have learned to a concrete example:

Exercise 4. Let us fix the initial data to be

$$x_0 = 1 \quad , \ u_0 = 0 \quad , \ y_0 = 1 \quad \text{and} \quad v_0 = \sqrt{3} \ .$$

Let $\mathbf{x}(t)$ be the unique solution of $\mathbf{x}''(t) = -\mathbf{x}(t)$ with $\mathbf{x}(0) = (x_0, y_0)$ and $\mathbf{x}'(0) = (u_0, v_0)$ for these values of the initial data.

(a) Find $R_2 = \max_{t\in\mathbb{R}} \|\mathbf{x}(t)\|$, and find all t for which $\|\mathbf{x}(t)\| = R_2$.

(b) Find $R_1 = \min_{t\in\mathbb{R}} \|\mathbf{x}(t)\|$, and find all t for which $\|\mathbf{x}(t)\| = R_1$.

(c) Sketch a plot of the ellipse traced out by $\mathbf{x}(t)$ at t ranges over $\mathbb{R}$.

SOLUTION For this initial data, $A = \begin{bmatrix} 1 & 0 \\ 1 & \sqrt{3} \end{bmatrix}$. Then $A^{-1} = \frac{1}{\sqrt{3}}\begin{bmatrix} \sqrt{3} & 0 \\ -1 & 1 \end{bmatrix}$, and

$$M = \frac{1}{3}\begin{bmatrix} 4 & -1 \\ -1 & 1 \end{bmatrix} \ .$$

We now compute that the characteristic polynomial of M is $t^2 - \frac{5}{3}t + \frac{1}{3}$, and then the eigenvalues of M are $\frac{1}{6}(5 \pm \sqrt{13})$. From here we find that the corresponding eigenvectors are (using the convention $\lambda_1 > \lambda_2$):

$$\mathbf{u}_1 := \frac{1}{\sqrt{2(13 - 3\sqrt{13})}}(2, 3 - \sqrt{13}) \quad \text{and} \quad \mathbf{u}_2 := \frac{1}{\sqrt{2(13 - 3\sqrt{13})}}(\sqrt{13} - 3, 2) \ .$$

We then compute

$$\mathbf{v}_1 := \frac{1}{\sqrt{(13-\sqrt{13})}}\left(\sqrt{6}\ ,\ -\frac{\sqrt{39}}{156}(13-\sqrt{13})\right)$$

and

$$\mathbf{v}_2 := \frac{1}{\sqrt{(13-\sqrt{13})}}\left(\frac{1}{2}\sqrt{6}(\sqrt{13}-3)\ ,\ -\frac{1}{\sqrt{2}}(\sqrt{13}-5)\right)\ .$$

We have seen that $\|\mathbf{x}(t)\|$ is maximal if and only if $(\cos t, \sin t) = \pm\mathbf{v}_2$ and is minimal if and only if $(\cos t, \sin t) = \pm\mathbf{v}_1$ We note that $\mathbf{v}_2 \approx (0.601, 0.799)$ and $\mathbf{v}_1 \approx (0.799, -0.601)$. Evidently, $\mathbf{v}_1$ lies in the first quadrant, and $-\mathbf{v}_1$ lies in the second quadrant. Applying arccos to the first entry of $\mathbf{v}_2$ gives a time $t_{\max}$ at which $\|\mathbf{x}(t)\|$ is maximal. Applying arccos to the first entry of $-\mathbf{v}_1$ gives a time $t_{\min}$ at which $\|\mathbf{x}(t)\|$ is miniimal.

Thus

$$t_{\max} = \text{arcos}\left(\frac{1}{2}\frac{\sqrt{6}(\sqrt{13}-3)}{\sqrt{(13-\sqrt{13})}}\right)\ ,$$

and the other times at which $\|\mathbf{x}(t)\|$ is maximal are $t_{\max} + k\pi$ where k is an integer. Likewise,

$$t_{\min} = \text{arcos}\left(\frac{\sqrt{6}}{\sqrt{(13-\sqrt{13})}}\right)\ ,$$

and the other times at which $\|\mathbf{x}(t)\|$ is maximal are $t_{\min} + k\pi$ where k is an integer.

In solving the concrete problem in Exercise 4, we have encountered some important constructions in Linear Algebra. Introducing the numbers σ_2 and σ_2 through

$$\sigma_1 = \frac{1}{\sqrt{\lambda_1}} \quad \text{and} \quad \sigma_2 = \frac{1}{\sqrt{\lambda_2}}\ ,$$

we can rewrite (0.9) as

$$A\mathbf{v}_1 = \sigma_1\mathbf{u}_1 \quad \text{and} \quad A\mathbf{v}_2 = \sigma_2\mathbf{u}_2 \tag{0.10}$$

where $\{\mathbf{u}_1, \mathbf{u}_2\}$ and $\{\mathbf{v}_1, \mathbf{v}_2\}$ are both orthonormal bases of $\mathbb{R}^2$.

Since any vector $\mathbf{x} \in \mathbb{R}^2$ can be written as

$$\mathbf{x} = (\mathbf{x}\cdot\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x}\cdot\mathbf{v}_2)\mathbf{v}_2\ ,$$

applying A to both sides, and using (0.10), we obtain

$$A\mathbf{x} = \sigma_1(\mathbf{x}\cdot\mathbf{v}_1)\mathbf{u}_1 + \sigma_2(\mathbf{x}\cdot\mathbf{v}_2)\mathbf{u}_2\ . \tag{0.11}$$

It turns out that this gives a factorization of the general invertible 2×2 matrix into a product of simple matrices. This is the *singular value decomposition.*

Exercise 5.

Introduce the matrices $V = [\mathbf{v}_1, \mathbf{v}_2]$, $S = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$ and $U = [\mathbf{u}_1, \mathbf{u}_2]$. (That is, the columns of V are $\mathbf{v}_1$ and $\mathbf{v}_2$, and the columns of U are $\mathbf{u}_1$ and $\mathbf{u}_2$.) Show that (0.11) is equivalent to

$$A\mathbf{x} = USV^T\mathbf{x}$$

all $\mathbf{x}$, and thus that the matrix A can be factored as the product of the three matrices U, S and V:

$$A = USV^T \ .$$

SOLUTION By the row formulation of matrix-vector multiplication, $V^T\mathbf{x} = (\mathbf{v}_1 \cdot \mathbf{x}, \mathbf{v}_2 \cdot \mathbf{x})$. Then $SV^T\mathbf{x} = (\sigma_1\mathbf{v}_1 \cdot \mathbf{x}, \sigma_2\mathbf{v}_2 \cdot \mathbf{x})$. Finally then

$$USV^T\mathbf{x} = \sigma_1(\mathbf{x} \cdot \mathbf{v}_1)\mathbf{u}_1 + \sigma_2(\mathbf{x} \cdot \mathbf{v}_2)\mathbf{u}_2 = A\mathbf{x} \ .$$

Recall that a square matrix with orthonormal columns is an *orthogonal matrix*, and that the inverse of an orthogonal matrix is its transpose. Hence U and V are orthogonal matrices, and their inverses are their transposes. The only 2×2 orthogonal matrices are rotations and reflections, so U and V both represent simple geometric transformations. The matrix S is even simpler: It represents a *scaling transformation*: It simply rescales the x and y coordinates by the factors σ_1 and σ_2 respectively.

Thus, every invertible linear transformation of $\mathbb{R}^2$ into itself is the product of an orthogonal transformation, followed by a scaling transformation, followed by another orthogonal transformation.

Here is one consequence of this important fact. Orthogonal transformations – rotations and reflections – do not affect area. Hence the effect on area of the linear transformation given by A comes entirely from the scale transformation in the middle. Thus, if R is a region in the plane, and $A(R)$ is its image under A the area of $A(R)$ is $\sigma_1\sigma_2$ times the area of R, and it is easy to see that $\sigma_1\sigma_2 = |\det(A)|$. All of this generalizes to matrices of arbitrary size, but in this workshop we will stay with 2 dimensions, and thoroughly deal with this case.

The final problem is to compute a singular value decomposition. Now that we know that such a decomposition exists, we can efficiently find it as follows. Starting from $A = USV^T$, we write

$$A^TA = (USV^T)^TUSV^T = VSU^TUSV^T = VS^2V^T$$

where we have used the fact that for any matrix product XY, $(XY)^T = Y^TX^T$, and the fact that since U^T is the inverse of U, since U is orthogonal. Then since V^T is the inverse of V, again, since V is orthogonal,

$$V^T(A^TA)V = V^{-1}(A^TA)V = S^2 \ .$$

Thus, V *diagonalizes the matrix* A^TA. This means the columns of V are an orthonormal basis of eigenvectors of A^TA, and the corresponding eigenvalues are the squares of the diagonal entries of S. So to find $\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\{\sigma_1, \sigma_2\}$, diagonalize A^TA. Finally, to find $\{\mathbf{u}_1, \mathbf{u}_2\}$, use (0.10).

Exercise 6. Let $A = \begin{bmatrix} 11 & -5 \\ 2 & -10 \end{bmatrix}$. Compute a singular value decomposition of A. That is, find orthogonal matrices U and V, and a diagonal matrix S with positive entries (these are the singular values) so that

$$A = USV^T \ .$$

SOLUTION

$$U = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}$$

$$S = 5\sqrt{2}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$V^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

All of the terms in the singular value decomposition have a direct meaning for the solution of the differential equation $\mathbf{x}'' = -\mathbf{x}$ when the matrix A is given by the initial data, as explained above. This is one way to think about where the singular value decomposition "comes from".