Lecture 3.31. Some applications of symmetric and orthogonal matrices

April 10, 2020

These classes are omnipresent in physics, as well as their complex counterparts, Hermitian and unitary matrices.

1. Mechanics and electricity.

Consider the system consisting of two masses which can slide on a horizontal rod, connected with three springs: the middle spring connects the masses, and two side springs connect the masses to fixed walls on the left and right hands of the segment. See Figure 5.3 on p. 275 of the book.

Suppose that the masses are m_1 and m_2, (m_1 on the left hand side) and the spring constants are k_0, k_1, k_2. Suppose that the masses move horizontally, without friction, so that the only forces acting on each mass come from the springs which obey the Hooke law. There is some equilibrium position (when nothing moves and all forces are balanced). Let $x_1(t)$ and $x_2(t)$ be small deviations of the masses from their equilibrium positions at time t. Then Newton's law and Hooke's law give the following system of differential equations:

$$m_1 \frac{d^2 x_1}{dt^2} = -k_0 x_1 + k_1(x_2 - x_1), \tag{1}$$

$$m_2 \frac{d^2 x_2}{dt^2} = -k_1(x_2 - x_1) - k_2 x_2. \tag{2}$$

Think carefully why the signs are as written! If $x_1 > 0$, the spring k_0 is stretched, so it tends to decrease x_1. If $x_2 > x_1$ then string k_1 is stretched, it tends to increase x_1. Similarly with the second equation. Introducing the vector $\mathbf{x} = (x_1(t), x_2(t))^T$ and the corresponding matrices, we write this as

$$\begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \mathbf{x}'' = \begin{pmatrix} -(k_0 + k_1) & k_1 \\ k_1 & -(k_1 + k_2) \end{pmatrix} \mathbf{x}. \tag{3}$$

This is a *second order* linear homogeneous differential equation with constant coefficients.

It describes *small oscillations* of a mechanical system. We could consider more masses and more springs, and obtain a system of larger size.

What I want to emphasize at this moment is that the matrix in the RHS is **symmetric**. This is not an accident, but it has a deep meaning. Please look at the equations and try to understand where this symmetry comes from. In the first equation, the term $k_1 x_2$ can be interpreted as the force with which m_2 acts on m_1. In the second equation, the force $-k_1 x_1$ can be interpreted as the force with which m_1 acts on m_2. These forces must be equal and have opposite directions according to the Newton's Third Law. Thus

The symmetry of the matrix in the right hand side of (3) is the direct consequence of the Third Law of Newton,

and since this is a fundamental law of nature, *any* system of this kind, describing small oscillations of a mechanical system near the equilibrium is expected to involve a symmetric matrix describing forces.

This is one reason why symmetric matrices are important in physics.

Let me mention that the same kind of differential equations describe also electric oscillations. Mathematically small oscillations of mechanical and electrical systems are the same thing. One simply has to substitute inductance for mass and reciprocal capacity for the spring constant. Then $x(t)$ can be interpreted either as the current or voltage at time t.

Speaking of solution, second order linear differential equations can be solved in the same way as first order ones: the method is simple: *plug* $\mathbf{x}(t) = e^{rt}\mathbf{v}$ *where* r *is a number and* $\mathbf{v}$ *some constant vector.* Suppose for simplicity that both masses are equal and set them to 1. then the system is

$$\mathbf{x}'' = K\mathbf{x}, \quad K = \begin{pmatrix} -k_0 - k_1 & k_1 \\ k_1 & -k_1 - k_2 \end{pmatrix}.$$

Plugging $\mathbf{x}(t) = s^{rt}\mathbf{v}$ and dividing on e^{rt} we obtain:

$$K\mathbf{v} = r^2\mathbf{v},$$

so r^2 is an eigenvalue, and $\mathbf{v}$ is an eigenvector.

The characteristic equation is

$$\lambda^2 + (k_0 + 2k_1 + k_2)\lambda + (k_0 + k_1)(k_1 + k_2) - k_1^2.$$

The discriminant

$$(k_0 + 2k_2 + k_2)^2 - 4(k_0 + k_1)(k_1 + k_2) + 4k_2^2 = (k_0 - k_2)^2 + 4k_2^2 > 0,$$

so the roots are real, their sum $-(k_0+2k_1+k_2)$ is negative, and their product $(k_0+k_1)(k_1+k_2)-k_1^2 = k_0k_1+k_0k_2+k_1k_2 > 0$ is positive. Therefore both roots are *negative*. This means that $r^2 < 0$, so we obtained four pure imaginary values for r; they form two complex conjugate pairs, and we denote them by $\pm i\omega_j, \quad j = 1, 2.$

Therefore, solutions will be oscillatory, with two oscillation frequencies $\omega_j/(2\pi)$. In example on p. 274-275, Strang takes some specific numbers and finds eigenvectors. Eigenvectors have a remarkable property: the one corresponding to the smaller frequency has all coordinates of the same sign, while the one for the larger frequency has coordinates of the opposite sign.

This means that oscillations of our system can be in two different *modes*. The low frequency mode corresponds to masses m_1, m_2 moving in the same direction, while for oscillations with higher frequency mode they move in the opposite directions.

A generic oscillation (with generic initial conditions) is a mixture (superposition) of these two modes.

By the way, the matrix K in this example is the "Strang's favorite matrix", and this example explains its origin.

The handout "String with beads" contains a remarkable complete solution for any number of equal masses and springs of equal elasticity.

2. Rotations in 3 space.

I recall that a rotation corresponds to an orthogonal matrix with determinant 1.

The characteristic equation of a 3×3 real matrix with determinant 1 has the form

$$p(\lambda) = -\lambda^3 + c_2\lambda^2 + c_1\lambda + 1 = 0.$$

So $p(0) = 1$ while $p(\lambda) \to -\infty$ as $\lambda \to +\infty$. This implies that there is a positive root λ_1. Since all eigenvalues of an orthogonal matrix have absolute value 1, we must have $\lambda_1 = 1$. Any eigenvector $\mathbf{v}$ corresponding to this eigenvalue will have the property $A\mathbf{v} = \mathbf{v}$, in other words, it is not moved by rotation. Such vectors form a 1-dimensional subspace (a line) which is called the *axis of rotation*. So we obtain

Euler's Theorem. *Every rotation in* $\mathbf{R}^3$ *has an axis.*

Notice that for $n = 2$ a rotation other than identity displaces every non-zero vector; there is no non-zero vector such that $A\mathbf{v} = \mathbf{v}$. Rotations in dimension 2 have no axis.

Exercise 1. *Give an example of rotation in dimension* 4 *which has no axis. That is find an orthogonal* 4×4 *matrix with determinant* 1 *for which* 1 *is not an eigenvalue.*

Exercise 2. *In which dimensions we have the property that every rotation has an axis? Hint: look where dimension* 3 *was used in the proof of Euler's theorem.*

Since the matrix of rotation is real, the other two eigenvalues are complex conjugate, or real. If they are real they are either both 1, in which case the rotation is just the identity, or both -1, for example

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Since the matrix of rotation is orthogonal, it preserves angles between vectors, $\mathrm{angle}(\mathbf{x}_1, \mathbf{x}_2) = \mathrm{angle}(A\mathbf{x}_1, A\mathbf{x}_2)$, in particular the orthogonal complement to the axis is mapped to itself. This orthogonal complement is a plane and rotation preserves this plane. Since the restriction to this plane is also a rotation, we can introduce orthogonal coordinates in this plane, and in these coordinates the matrix will be the same as of two-dimensional rotation. So any rotation is characterized by the *axis and the angle of rotation.*

For example, rotations whose axis is the x_3-axis have matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

Similarly one can write rotation matrices about x_2 and x_3 axes.

Again, to encode a rotation we need the axis and the angle. The axis can be defined by any non-zero vector in it, its length and direction does not matter. For the angle, there is also some ambiguity. Let us agree to choose the angle between $-\pi$ and π, and remember that angles $\pm\pi$ give the *same* rotation.

To describe the axis, let us choose the vector on it of length $|\theta|$. Then a single vector will encode the rotation: its length is the absolute value of the angle, and direction is the direction of the axis. There are still two vectors to choose from, so we need some rule. The common rule is the

Right hand rule: if you look in the direction of the vector, the positive direction of rotation is clockwise.

Or, equivalently, it is the direction of rotation of a right screw when you drive it in the direction of the vector.

These rules assign exactly one rotation to every vector of length $< \pi$, but still two opposite vectors of length π correspond to the same rotation.

In other words, all rotations are in correspondence to points of the closed ball of radius π centered at the origin; to each vector in this ball corresponds a unique rotation; to each rotation by an angle other than $\pm\pi$ corresponds a unique vector, but to rotations by π two vectors correspond. The endpoints of these two vectors lie at the diametrally opposite points on the surface of the ball.

We denote by $R(\mathbf{a})$ the rotation corresponding to a vector $\mathbf{a}$, $\|\mathbf{a}\| \leq \pi$.

The next question: how to find the matrix of $R(\mathbf{a})$ for a given a?

One method of doing this is evident:

For given a, find a basis in the orthogonal complement of $\mathbf{a}$, orthognalize this basis to obtain an orthonormal basis $(v_1, v_2, a/\|\mathbf{a}\|)$. In this basis the rotation will have matrix (4). Then transform to the standard basis.

I will explain a more elegant and illuminating method, using the exponential. In the previous lecture, it was proved that for every skew-symmetric matrix A, e^A is a rotation, and I hinted that in fact every rotation has this form. This is what we are going to prove now.

Any skew-symmetric 3×3 matrix can be written as

$$A = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}.$$

Let us investigate $\exp(A)$. To each vector $\mathbf{a} = (a_1, a_2, a_3)^T \in \mathbf{R}^3$ corresponds auch a matrix.

The characteristic polynomial is

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & -a_3 & a_2 \\ a_3 & -\lambda & -a_1 \\ -a_2 & a_1 & -\lambda \end{vmatrix} = -\lambda^3 - (a_1^2 + a_2^2 + a_3^2)\lambda,$$

so eigenvalues are $0, \pm i\sqrt{a_1^2 + a_2^2 + a_3^2} = \pm i\|\mathbf{a}\|$, where $\mathbf{a} = (a_1, a_2, a_3)^T$. The eigenvectors corresponding to $\lambda_1 = 0$ must satisfy

$$A\mathbf{v} = 0,$$

and we see that $\mathbf{a}$ is an eigenvector!

By the Spectral mapping theorem, e^A will have the same eigenvector $\mathbf{a}$ with eigenvalue $e^0 = 1$, so the direction of rotation is $\mathbf{a}$, and the angle of rotation is $\pm\|a\|$. In other words, $e^A = R(\mathbf{a})$ and we showed that rotation encoded by a vector $\mathbf{a}$ is exactly e^A where A is as above.