Lecture 4.14. Simultaneous diagonalization of two quadratic forms and a generalized eigenvalue problem

A. Eremenko

April 10, 2020

In the previous lectures we have seen that every symmetric matrix can be brought to a diagonal form by a change of the basis. This change of the basis is not unique. It turns out that in an important special case we can bring *two matrices* to a diagonal form by the *same* change of the basis.

Theorem. *Let A, M be two real symmetric matrices of the same size, and let M be positive definite. Then there exists a non-singular matrix C such that*

$$C^T MC = I, \tag{1}$$

and

$$C^T AC = \Lambda, \tag{2}$$

where Λ is s real a diagonal matrix.

Proof. We have

$$M = R^T R, \tag{3}$$

with some non-singular matrix R. Then the matrix

$$(R^{-1})^T AR^{-1}$$

is symmetric, so there exists an orthogonal matrix B such that

$$B^{-1}(R^{-1})^T AR^{-1}B = \Lambda. \tag{4}$$

Set

$$C = R^{-1}B. \tag{5}$$

Then $C^T = B^T(R^{-1})^T = B^{-1}(R^{-1})^T$, where we used that B is orthogonal. So (4) is the same as (2). To check (1) we use $B^T = B^{-1}$ and $(R^{-1})^T = (R^T)^{-1}$ and obtain

$$C^T MC = B^{-1}(R^T)^{-1}R^T RR^{-1}B = I.$$

This proves Theorem 1.

Generalized eigenvalue problem

Next we will show that the entries λ_j of the diagonal matrix Λ in this theorem are *generalized eigenvalues* of A with respect to M:

$$Ax = \lambda Mx, \quad x \neq 0. \tag{6}$$

They can be determined from the *generalized characteristic equation*

$$\det(A - \lambda M) = 0,$$

and the *generalized eigenvectors* are the columns of C from (1), (2).

We obtain from Theorem 1 and from its proof:

Corollary. *Let A, M be symmetric matrices of the same size, and let M be positive definite. Then all generalized eigenvalues (6) are real, and there is a basis of the whole space which consists of generalized eigenvectors.*

Proof. We refer to the proof of Theorem 1. Matrix $(R^{-1})^T AR^{-1}$ is symmetric, therefore all its eigenvalues are real and the eigenvectors form a basis. These eigenvectors are columns of B. If v_j is an eigenvector of $(R^{-1})^T AR^{-1}$ with eigenvalue λ_j then

$$(R^{-1})^T AR^{-1}v_j = \lambda_j v_j.$$

Multiplying on R^T and setting $v_j = Ru_j$ we obtain

$$Au_j = \lambda_j R^T Ru_j = \lambda_j R^T Ru_j = \lambda_j Mu_j.$$

As R is non-singular, the u_j form a basis of the space.

Since $B = RC$, this basis is nothing but the columns of C of Theorem 1. So the simultaneous diagonalization of two matrices is not more difficult than diagonalization of one matrix: solve the generalized characteristic equation and find generalized eigenvectors.

Notice that

$$u_i^T M u_j = u_i R^T R u_j = (Ru_i)^T (Ru_j) = v_i^T v_j = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

This means that u_j, the eigenvectors of (6) are orthonormal with respect to the dot product defined by

$$(x, y)_M = x^T M y,$$

and our matrix R transforms this dot product to the standard dot product:

$$(x, y)_M = x^T M y = x^T R^T R y = (Rx, Ry).$$

Geometric interpretation

I recall the geometric interpretation of a positive definite quadratic form. It defines an ellipsoid:

$$\{\mathbf{x} : \mathbf{x}^T A \mathbf{x} = 1\}.$$

An ellipsoid can be rotated so that its principal axes become the coordinate axes. An ellipsoid can be also stretched along the principal axes so that it will become a sphere.

Now suppose we have two positive definite matrices, which define two ellipsoids, say E_1 and E_2. We can perform a stretching along the axes of the first ellipsoid E_1 to make it into a sphere. Under this stretching, the second ellipsoid will become some new ellipsoid E_3. Then we can perform a rotation which makes the principal axes of E_3 coincide with coordinate axes, and this rotation will not change the sphere.

This means that a composition of the stretching and the rotation transforms E_1 into a sphere and brings E_2 to principal axes. This is the geometric meaning of our theorem.

Applications to mechanics.

Let me recall the example from Lecture 3.31 of two masses connected by springs and moving on an interval without friction. The system is described by the differential equation

$$M\mathbf{x}'' = -K\mathbf{x}, \tag{7}$$

where

$$M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} k_0 + k_1 & -k_1 \\ -k_1 & k_1 + k_2 \end{pmatrix}.$$

In this example, both matrices are positive definite, and this has a physical meaning: M is positive definite because mass is always positive, and K is positive definite because of the Hooke's Law. To check the positivity of K we use the NW minor criterion:

$$k_0 + k_1 > 0$$

and

$$\det K = (k_0 + k_1)(k_1 + k_2) - k_1^2 = k_0 k_1 + k_0 k_2 + k_1 k_2 > 0.$$

Let us apply our theorem and find a non-singular matrix C such that

$$C^T M C = I, \quad C^T K C = \Lambda.$$

Let us change the variable $\mathbf{x} = C\mathbf{y}$. Introducing this to our equation (7) and multiplying both sides on C^T we obtain

$$C^T M C \mathbf{y}'' = -C^T K C \mathbf{y},$$

or

$$\mathbf{y}'' = -\Lambda \mathbf{y}.$$

So our equation decouples: it is a trivial system of scalar equations

$$\begin{aligned} y_1'' &= -\lambda_1 \mathbf{y}_1 \\ y_2'' &= -\lambda_2 \mathbf{y}_2. \end{aligned}$$

Since both λ_j are positive we obtain solutions

$$\begin{aligned} y_1(t) &= c_1 \cos \omega_1 t + c_2 \sin \omega_1 t, \\ y_2(t) &= c_3 \cos \omega_2 t + c_4 \sin \omega_2 t, \end{aligned}$$

where $\omega_j = \sqrt{\lambda_j}$. So in the new coordinates, the motion is a superposition of two harmonic oscillations, whose frequencies are square roots of the generalized eigenvalues.

This example is quite general: the matrices can be of any size, and the system describes any kind of small oscillations in mechanics or electricity.

Example. Let us take

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

The characteristic equation

$$\det(K - \lambda M) = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-2\lambda \end{vmatrix} = 2\lambda^2 - 6\lambda + 3.$$

Its roots are

$$\lambda_1 = \frac{3-\sqrt{3}}{2}, \quad \lambda_2 = \frac{3+\sqrt{3}}{2}.$$

Notice that both roots are positive. To the first root corresponds an eigenvector $\mathbf{v}_1 = (\sqrt{3}-1, 1)^T$ and to the second root the eigenvector $\mathbf{v}_2 = (\sqrt{3}+1, -1)^T$. They describe the *modes* of oscillations. The eigenvectors that we found are M-orthogonal:

$$\mathbf{v}_1^T M \mathbf{v} = (\sqrt{3}-1, 1) \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \sqrt{3}+1 \\ -1 \end{pmatrix} = 0.$$

Notice that the eigenvector corresponding to the smaller eigenvalue has no sign change. This is a special case of quite general phenomenon: describing such oscillations.

General mechanical systems

The following is just an example of further applications. It uses some notions of mechanics, which are briefly explained. The goal is to show the importance and generality of the previous example. No knowledge of mechanics will be required on the exam.

Newton's form of equations of motion $ma = F$ is not always convenient, especially when one deals with curvilinear coordinates. A generalization was proposed by Lagrange. We consider a system of points whose position is completely determined by some *generalized coordinates* $\mathbf{q} = (q_1, \ldots, q_n)$. For example, for one free point in space we have three coordinates $(q_1, q_2, q_3) = (x_1, x_2, x_3)$. Or $\mathbf{q}$ may be cylindrical, or spherical coordinates. For m free points in space we need $n = 3m$ coordinates. For a pendulum oscillating in a vertical plane, we need one coordinate, for example the angle of deviation of this pendulum from the vertical is a convenient coordinate.

As the system moves, coordinates are functions of time $q_j(t)$. Their derivatives are called *generalized velocities*, $\dot{q} = dq/dt$. Derivatives with respect to time are usually denoted by dots over letters in mechanics, to distinguish them from other derivatives. To obtain the true velocity vector of a point $\mathbf{x}_k \in \mathbf{R}^3$, one has to write $\mathbf{x}_k = f_k(\mathbf{q})$, rectangular coordinates as functions of generalized coordinates, and differentiate:

$$\dot{\mathbf{x}}_k = \sum_{j=1}^{n} \frac{\partial f_k}{\partial q_j} \dot{q}_j,$$

and the kinetic energy is

$$T_k = m\|\dot{\mathbf{x}}_k\|^2/2 = \sum b_{k,i,j}(\mathbf{q})\dot{q}_i\dot{q}_j, \tag{8}$$

where $b_{k,i,j}$ are some functions of $\mathbf{q}$. The total kinetic energy T of the system is the sum of such expressions over all points $\mathbf{x}_k$, $T = \sum T_k$. The important fact is that

Kinetic energy is a positive definite quadratic form of generalized velocities, with coefficients depending on the generalized coordinates.

It is positive definite because the LHS of (8) is non-negative and the sum of such expressions is positive, if at least one point actually moves.

Now we assume that *vector of forces is the gradient of some function $-U$ of generalized coordinates*:

$$F = -\operatorname{grad} U = -\left(\frac{\partial U}{\partial q_1}, \ldots, \frac{\partial U}{\partial q_n}\right). \tag{9}$$

This function U is called the potential energy or simply the *potential.*

Following Lagrange's recipe, we form the following function of generalized coordinates and velocities:

$$L = T - U,$$

the difference between kinetic and potential energy. This function is called the *Lagrangian* of the system. The equations of motion in the form of Lagrange are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{\partial L}{\partial q_j}, \quad 1 \leq j \leq n. \tag{10}$$

The advantage of this formulation is that unlike for Newton's equations *arbitrary* curvilinear coordinate system can be used.

To see that these equations indeed generalize Newton's equations, consider a free point with coordinate $\mathbf{x} = (x_1, x_2, x_3)$ and mass m moving in the field of force with potential U. Then the kinetic energy is

$$T = \frac{m}{2}\left(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2\right),$$

and the Lagrangian is $L = T - U$. So equations (10) become

$$\frac{d}{dt}(m\dot{x}_j) = -\frac{\partial U}{\partial x_j} = F_j(x_1, x_2, x_3),$$

where F_j is the j-th component of the force.

Equations of motion are usually non-linear and cannot be solved.

One of the most common methods of dealing with them is *linearization*, that is approximation of non-linear equations by linear ones. The simplest case is the linearization near an equilibrium. An equilibrium is a point $\mathbf{q}^0$ such that the system in this state does not move. This means that equations (10) are satisfied by $\mathbf{q}(t) \equiv \mathbf{q}^0$.

Theorem 2. *A point $\mathbf{q}^0$ is an equilibrium if and only if it is a critical point of the potential energy U.*

Proof. Let us write (10) as

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} = \frac{\partial T}{\partial q_j} - \frac{\partial U}{\partial q_j}.$$

If $\mathbf{q}(t) \equiv \mathbf{q}^0$ is a solution, then $\dot{\mathbf{q}} = 0$ and thus $\partial T/\partial \dot{q}_j = 0$, and $\partial T/\partial q_j = 0$ for all j. So $\partial U/\partial q_j = 0$.

For the linearization we assume without loss of generality that $\mathbf{q}^0 = 0$, and that both T and U are *analytic* functions of $\mathbf{q}$ and $\dot{\mathbf{q}}$. This means that they have convergent series expansions

$$T(q, \dot{q}) = T_0 + T_1(q, \dot{q}) + T_2(q, \dot{q}) + \dots,$$

where T_k are homogeneous polynomials of the variables $q_j, \dot{q}_j$. Similar expansion holds for U. When we differentiate a homogeneous polynomial, its degree decreases by 1, so to obtain *linear* equations in (10) both T and U have to be of degree 2. As T is of degree 2 in the variables $\dot{q}$, it must be independent of q. In U, the terms of the first degree vanish by Theorem 2, and the constant term disappears after differentiation.

Thus

The Lagrangian of the linearized systems at an equilibrium point 0 *is obtained by setting* $\mathbf{q} = 0$ *in (8) and keeping only quadratic terms in U, in other words, this Lagrangian has the form*

$$L = \sum_{i,j} m_{i,j}\dot{q}_i\dot{q}_j - \sum_{i,j} a_{i,j}q_iq_j, \tag{11}$$

the difference of two quadratic forms with matrices M and A such that M is positive definite.

To write the Lagrange equations with Lagrangian (11) we need the differentiation formula

$$\frac{d}{d\mathbf{x}}(\mathbf{x}^T A\mathbf{x}) = A\mathbf{x}.$$

Here $d/d\mathbf{x}$ is the column $(d/dx_1, \ldots, d/dx_n)^T$. So the equation of motion with Lagrangian (11) is

$$\frac{d}{dt}M\dot{\mathbf{x}} = -A\mathbf{x}. \tag{12}$$

Now we can use the theorem on simultaneous diagonalization. We can apply it directly to (11) to conclude that there are new coordinates $\mathbf{y}$, $\mathbf{x} = C\mathbf{y}$, such that

$$L = \dot{\mathbf{y}}^T\dot{\mathbf{y}} - \mathbf{y}^T\Lambda\mathbf{y},$$

so the equations of motion *decouple* and become

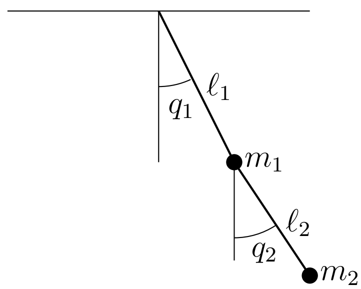
$$\ddot{y}_j = -\lambda_j y_j. \tag{13}$$

Or alternatively we can apply the Corollary to the linear equation (12) and find a basis u_j of generalized eigenvectors. If y_j are coordinates with respect to this basis, we obtain (13) again. Stated in words this means that *the linearized equation of small oscillations always decouples and becomes (13) after a change of coordinates.*

Notice that if the matrix A is positive definite, then all $\lambda_j > 0$ and solutions have the form $y_j(t) = c_j e^{\pm i\omega_j}$, where $\omega_j = \sqrt{\lambda_j}$ the system is stable and solutions oscillate with frequencies ω_j.

Example. Double pendulum.

The configuration is shown in the figure. Let us choose the angles between the two rods and the vertical direction as generalized coordinates q_1, q_2. Angles are measured from the downward vertical direction, counterclockwise, as shown in the picture.



Then the kinetic energy of the mass m_1 is

$$T_1 = \frac{m_1}{2}\ell_1^2\dot{q}_1^2,$$

and the kinetic energy of the second mass is

$$T_2 = \frac{m_2}{2}\left(\ell_1^2\dot{q}_1^2 + \ell_2^2\dot{q}_2^2 + 2\ell_1\ell_2\cos(q_2-q_1)\dot{q}_1\dot{q}_2\right).$$

Potential energy of the system is

$$U = -(m_1+m_2)g\ell_1\cos q_1 - m_2 g\ell_2\cos q_2.$$

Thus $T = T_1 + T_2$ and

$$\begin{aligned} L = T - U &= \frac{m_1+m_2}{2}\ell_1^2\dot{q}_1^2 + \frac{m_2}{2}\ell_2^2\dot{q}_2^2 + m_2\ell_1\ell_2\cos(q_2-q_1)\dot{q}_1\dot{q}_2 \\ &+ (m_1+m_2)g\ell_1\cos q_1 + m_2 g\ell_2\cos q_2. \end{aligned}$$

The equations of motion are non-linear and difficult to solve, so we linearize them near the equilibrium $(q_1, q_2) = (0, 0)$. (There are four equilibria in our system). Linearization in this case means that we replace the cosine in the kinetic energy by 1 and the cosines in potential according to the formula $\cos x \approx 1 - x^2/2$, because we want to keep only second degree terms in the Lagrangian. The constant term in Potential energy can be omitted.

Thus the Lagrangian of the linearized system is

$$\begin{aligned} L^* &= \frac{m_1+m_2}{2}\ell_1^2\dot{q}_1^2 + \frac{m_2}{2}\ell_2^2\dot{q}_2^2 + m_2\ell_1\ell_2\dot{q}_1\dot{q}_2 \\ &- \frac{m_1+m_2}{2}g\ell_1 q_1^2 - \frac{m_2}{2}g\ell_2 q_2^2. \end{aligned}$$

and the linearized equation of motion is

$$\begin{pmatrix} (m_1+m_2)\ell_1^2 & m_2\ell_1\ell_2 \\ m_2\ell_1\ell_2 & m_2\ell_2^2 \end{pmatrix} \begin{pmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{pmatrix} = - \begin{pmatrix} (m_1+m_2)g\ell_1 & 0 \\ 0 & m_2 g\ell_2 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix},$$

which we write as

$$M\ddot{q} = -Aq.$$

It is easy to check directly that both M and A are positive definite. Notice that M is not diagonal in this example.