5. Classification of second-order equations

There are 2 general methods for classifying higher-order partial differential equations. One is *very* general (applying even to some nonlinear equations), and seems to have been motivated by the success of the theory of first-order PDEs. In this method one rewrites the higher order PDE as a *system* of *first-order* PDEs and attempts to generalize the method of characteristics to that system. This turns out to be possible only for a restricted (but important) class of PDEs called *hyperbolic*.

The other classification method applies only to linear, second-order equations. This lecture covers this second method; we postpone the first method until the last lecture.

Consider the general linear, second-order PDE in the form

$$(1) \qquad \sum_{ij} A_{ij} \frac{\partial^2 \theta}{\partial x_i \partial x_j} + \sum_i B_i \frac{\partial \theta}{\partial x_i} + C\theta = F(\mathbf{x})$$

where θ is a function of the n variables $\mathbf{x} \equiv (x_1, x_2, \ldots, x_n)$. We assume that all the coefficients A_{ij}, B_i and C are constants. With no loss in generality, we assume that A_{ij} is symmetric.

We will show that, because A_{ij} is symmetric, it is possible to transform to new coordinates in which (1) takes the diagonal form

$$(2) \qquad \sum_i A_i \frac{\partial^2 \theta}{\partial x_i^2} + \sum_i B_i \frac{\partial \theta}{\partial x_i} + C\theta = F(\mathbf{x})$$

and each A_i takes the value $+1$, -1 or 0. For example, the general 2-dimensional equation

$$(3) \qquad A_{11}\theta_{xx} + 2A_{12}\theta_{xy} + A_{22}\theta_{yy} + l.o.d. = F$$

(where *l.o.d.* means "lower-order derivatives") can always be transformed into one of the following forms:

$$(4) \qquad \begin{aligned} &\theta_{xx} + \theta_{yy} + l.o.d. = F \\ &\theta_{xx} - \theta_{yy} + l.o.d. = F \\ &\theta_{xx} \pm \theta_{y} + l.o.d. = F \end{aligned}$$

plus various permutations of these, such as $\theta_{yy} \pm \theta_x + l.o.d. = F$. In (4) we write only the highest-derivative terms in each variable, and we do not consider the case $\theta_x + \theta_y + \cdots$ because it is a first-order equation.

It turns out that some of the most important properties of such equations (such as the appropriate form of boundary conditions) depend only on the highest-derivative terms. Thus there is some point in considering the cases (4) in their "purest forms":

$\theta_{xx} + \theta_{yy} = F$ Poisson's equation

(5) $\theta_{tt} - \theta_{xx} = F$ wave equation

$\theta_t - \theta_{xx} = F$ heat equation

(where the notation hints at the typical physical meaning).

To prove our claim we must first prove a theorem about symmetric A_{ij}. For future purposes, we shall prove more than is actually needed for this lecture.

Theorem. [The spectral decomposition theorem.] If the matrix **A** is Hermitian (meaning that $A_{ij} = A_{ji}*$ where * denotes the complex conjugate), then all the eigenvalues of **A** are real, and all the eigenvectors are, or can be made, orthogonal.

Note that real symmetric **A**, like that in (1), are a class of Hermitian matrices.

Before proving this theorem, we note that the problem of diagonalizing (1) is closely related to the problem of diagonalizing the *quadratic form*

$$\sum_{ij} A_{ij} x_i x_j + \sum_i B_i x_i + C\theta . \tag{6}$$

In both problems the strategy is the same: choose the new coordinates $\bar{\mathbf{x}}$ to be the coefficients in the expansion

$$\mathbf{x} = \sum_i \bar{x}_i \mathbf{e}_i \tag{7}$$

of **x** in the *n* eigenvectors $\mathbf{e}_i$ of **A**. Then the first term in (6) becomes

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= \left(\sum_j \bar{x}_j \mathbf{e}_j \right)^T \mathbf{A} \left(\sum_i \bar{x}_i \mathbf{e}_i \right) \\ &= \sum_{i,j} \bar{x}_i \bar{x}_j \left(\mathbf{e}_j{}^T \mathbf{A} \mathbf{e}_i \right) = \sum_{i,j} \bar{x}_i \bar{x}_j \left(\mathbf{e}_j{}^T \lambda_i \mathbf{e}_i \right) = \sum_{i,j} \bar{x}_i \bar{x}_j \lambda_i \, \mathbf{e}_j \cdot \mathbf{e}_i = \sum_i \lambda_i \bar{x}_i^2 \end{aligned} \tag{8}$$

where λ_i is the eigenvalue corresponding to eigenvector $\mathbf{e}_i$. Note that we use the supposed orthogonality of the eigenvectors in 2 ways: first, in the assumption (7) that any **x** can be expanded in these eigenvectors, and, second, in the final step of (8). By an additional rescaling of $\bar{x}_i$, we may reduce (8) to simplest form, $\sum_i \bar{x}_i^2$.

The transformation of (1) to the diagonal form (2) proceeds in the same manner. But first we prove the theorem.

Sketch of the proof. Recall that the column vector **e** is an eigenvector of the $n \times n$ matrix **A** if $\mathbf{Ae} = \lambda \mathbf{e}$ where λ is the corresponding eigenvalue, a generally complex number. Thus the eigenvalues of **A** correspond to the roots of det $(\mathbf{A} - \lambda \mathbf{I}) = 0$. Let λ_i be the eigenvalue of **A** corresponding to eigenvector $\mathbf{e}_i$:

$$\mathbf{A}\,\mathbf{e}_i = \lambda_i \mathbf{e}_i . \tag{9}$$

Let $\mathbf{A}^+ \equiv \mathbf{A}^{T*}$ (transpose conjugate). Thus $\mathbf{A} = \mathbf{A}^+$ if **A** is Hermitian.

To prove that λ_i is real, take the transpose conjugate of (9) to get

(10) $\quad \mathbf{e}_i^{\,+}\mathbf{A}^+ = \lambda_i^{\,*}\mathbf{e}_i^{\,+}.$

(Note that $\mathbf{e}_i^{\,+}$ is a row vector.) Now multiply (9) on the left by $\mathbf{e}_i^{\,+}$, multiply (10) on the right by $\mathbf{e}_i$, and subtract, using the fact that $\mathbf{A} = \mathbf{A}^+$. The result is

(11) $\quad \left(\lambda_i - \lambda_i^{\,*}\right)\mathbf{e}_i^{\,+}\mathbf{e}_i = 0$

which proves that λ_i is real.

To prove that the $\mathbf{e}_i$ are orthogonal, consider any 2 eigenvalues λ_i and λ_j. Multiply (9) by $\mathbf{e}_j^{\,+}$, multiply

(12) $\quad \mathbf{e}_j^{\,+}\mathbf{A}^+ = \lambda_j\mathbf{e}_j^{\,+}$

(obtained by taking the Hermitian conjugate of (9)) by $\mathbf{e}_i$ on the right and subtract, again using the fact that **A** is Hermitian, to obtain

(13) $\quad \left(\lambda_i - \lambda_j\right)\mathbf{e}_j^{\,+}\mathbf{e}_i = 0.$

Thus if $\lambda_i \neq \lambda_j$, then

(14) $\quad \mathbf{e}_j^{\,+}\mathbf{e}_i = \mathbf{e}_j \cdot \mathbf{e}_i = 0$

and the two eigenvectors are orthogonal. This proof does not apply to the case $\lambda_i = \lambda_j$ of degenerate eigenvalues, which we postpone until later in this lecture.

Assume for the moment that the n eigenvalues are distinct. Then we have n orthogonal eigenvectors. These can easily be normalized; we henceforth assume that they are. The normalized eigenvectors can be used to diagonalize **A**. Consider the $n \times n$ matrix

(15) $\quad \mathbf{U} = \left(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\right)$

whose column vectors are the orthonormal eigenvectors of **A**. Then

(16) $\quad \mathbf{U}^+ = \mathbf{U}^{-1},$

that is,

(17) $\quad \mathbf{U}^+\mathbf{U} = \mathbf{I}$

which simply restates the fact that the $\mathbf{e}_i$ are orthonormal. Matrices **U** with the property (17) are called *unitary*. [Note that since **A** and λ_i are real, it is always possible to choose real eigenvectors $\mathbf{e}_i$. Then $\mathbf{U}$ is real and $\mathbf{U}^+ = \mathbf{U}^T$]. The matrix $\mathbf{U}$ defined by (15) diagonalizes $\mathbf{A}$ in the sense that

(18) $$\begin{aligned} &\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{U}^T\mathbf{A}\mathbf{U} = \mathbf{U}^T\left(\lambda_1\,\mathbf{e}_1, \lambda_2\,\mathbf{e}_2, \ldots, \lambda_n\,\mathbf{e}_n\right) \\ &= diag\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right) \end{aligned}$$

Now we return to the general equation (1) with second-order term

(19) $$\sum_{ij} A_{ij} \frac{\partial^2\theta}{\partial x_i \partial x_j}$$

Defining U_{ij} as above, we introduce the new variables

(20) $$\overline{x}_i = \sum_j U_{ij}{}^T x_j .$$

Thus $\overline{x}_i$ is the projection of $\mathbf{x}$ onto $\mathbf{e}_i$; (20) is equivalent to (7). Since

(21) $$\frac{\partial}{\partial x_i} = \sum_j \frac{\partial \overline{x}_j}{\partial x_i} \frac{\partial}{\partial \overline{x}_j} = \sum_j U_{ji}{}^T \frac{\partial}{\partial \overline{x}_j}$$

we have

(22) $$\sum_{ij} A_{ij} \frac{\partial^2\theta}{\partial x_i \partial x_j} = \sum_{ijkm} A_{ij} U_{ki}{}^T U_{mj}{}^T \frac{\partial^2\theta}{\partial \overline{x}_k \partial \overline{x}_m} = \sum_{km} \overline{A}_{km} \frac{\partial^2\theta}{\partial \overline{x}_k \partial \overline{x}_m}$$

where, in matrix notation,

(23) $$\overline{\mathbf{A}} = \mathbf{U}^T\mathbf{A}\mathbf{U} .$$

But, by (18), $\overline{\mathbf{A}}$ is the diagonal matrix with the eigenvalues of $\mathbf{A}$ on the diagonal. Thus

(24) $$\sum_{ij} A_{ij} \frac{\partial^2\theta}{\partial x_i \partial x_j} = \sum_i \lambda_i \frac{\partial^2\theta}{\partial \overline{x}_i^{\,2}} .$$

This is almost the simplest form. To obtain the simplest form, it is only necessary to redefine

(25) $$\overline{x}_i \leftarrow \overline{x}_i \,/\, \sqrt{|\lambda_i|}$$

(provided that $\lambda_i \neq 0$.)

In overall summary, to transform the second derivatives in (1) to the canonical form (2), we use the transformation

(26) $$\overline{\mathbf{x}} = \mathbf{U}^T\mathbf{x}$$

where **U** is the unitary matrix defined by (15). The components of $\bar{\mathbf{x}}$ are just the "amplitude" of each eigenvector's contribution to **x**. Since unitary transformations have the property that they preserve $\mathbf{x}\cdot\mathbf{x}$ (prove this!), the transformation (26) may be interpretted as a rotation of the coordinates in *n*-dimensional space.

How does this work in the case *n*=2? We let $a = A_{11}$, $b = A_{12}$, and $c = A_{22}$, so (1) takes the form

(27) $$a\theta_{xx} + 2b\theta_{xy} + c\theta_{yy} + \cdots = F$$

The eigenvalues of

(28) $$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

satisfy

(29) $$(a-\lambda)(c-\lambda) = b^2.$$

Thus

(30) $$\lambda = \frac{a+c}{2} \pm \frac{1}{2}\sqrt{(a-c)^2 + 4b^2} = \frac{a+c}{2} \pm \frac{1}{2}\sqrt{(a+c)^2 + 4(b^2 - ac)}$$

from which we see that λ is indeed real for any *a, b, c*. If

(31) $$b^2 - ac = 0 \qquad \text{(the parabolic case)}$$

it follows from (30) that at least one of the eigenvalues vanishes. This is the *parabolic case* and leads to $\theta_{xx} + \theta_y + \cdots$ or $\theta_{yy} + \theta_x + \cdots$

Similarly, the 2 eigenvalues have the same sign if

(32) $$b^2 - ac < 0 \qquad \text{(the elliptic case)}$$

This is the *elliptic case* and leads to $\theta_{xx} + \theta_{yy} + \cdots$

Finally, the 2 eigenvalues have opposite signs if

(33) $$b^2 - ac > 0 \qquad \text{(the hyperbolic case)}$$

This is the *hyperbolic case* and leads to $\theta_{xx} - \theta_{yy} + \cdots$ We have already mentioned the connection with quadratic forms. For the form $ax^2 + 2bxy + cy^2$, the same transformation produces a diagonalization, and the resulting equation describes a parabola, ellipse or hyperbola.

Quadratic forms lie at the heart of Riemannian geometry. There one seeks coordinates in which the differential arc length *ds*, which has the general form

(34) $$ds^2 = \sum_{ij} A_{ij} dx_i dx_j$$

simplifies to

(35) $$ds^2 = (d\bar{x}_1)^2 \pm (d\bar{x}_2)^2 \cdots \pm (d\bar{x}_n)^2$$

The new coordinates are called Cartesian coordinates. Examples include

Euclidean space $\quad ds^2 = (dx)^2 + (dy)^2 + (dz)^2$

Lorentzian spacetime $\quad ds^2 = c^2(dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$

If you have studied general relativity (where A_{ij} is always written as g_{ij}) then you know that global Cartesian coordinates exist only if the curvature tensor associated with g_{ij} vanishes.

This bears upon the generalization of (1) to the case of nonconstant coefficients $A_{ij}(\mathbf{x})$. Whereas (1) can always be transformed locally to the form (2), there is in general no globally valid transformation. For the case n=2 however, one can show that a global transformation to the form (2) is possible provided that the A_{ij} do not vary in such a way that the equation changes type (e.g. from elliptic to hyperbolic). For a careful discussion of this, see Garabedian, Chapter 2.

We must still complete the proof of the spectral decomposition theorem by showing that orthonormal eigenvectors may be found even in the case where some of the eigenvalues are equal.

Completion of the proof. Even if all the eigenvalues are equal, we can find at least one eigenvector $\mathbf{e}_1$. Then we define

(36) $$\mathbf{U}_1 = (\mathbf{e}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_n)$$

where $\hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_n$ are $n-1$ other vectors (not necessarily eigenvectors of $\mathbf{A}$) which are orthonormal to $\mathbf{e}_1$ and to each other. (It is always possible to find such a set!) By the orthonormality of its columns, $\mathbf{U}_1$ is unitary.

We use $\mathbf{U}_1$ to transform $\mathbf{A}$ into the form

(37) $$\mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{A}_2 & \\ 0 & & & \end{pmatrix}$$

where $\mathbf{A}_2$ is an $(n-1)\times(n-1)$ matrix. Then we do the same thing to $\mathbf{A}_2$. That is, we find the (unitary) $\mathbf{U}_2$ such that

(38) $$\mathbf{U}_2^{-1}\left(\mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1\right)\mathbf{U}_2 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & 0 & & & \\ & \vdots & & \mathbf{A}_3 & \\ 0 & 0 & & & \end{pmatrix}$$

and keep going. Each step corresponds to a rotation in the remaining coordinates. At the end we have

(39) $$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right)$$

where $\mathbf{U} = \mathbf{U}_1\mathbf{U}_2\cdots\mathbf{U}_n$ is unitary because the product of unitary matrices is unitary (prove this!). Eqn (39) implies that the columns of **U** are the sought-for orthonormal eigenvectors. (Each $\mathbf{U}_i$ has rank *n*, hence so does **U**.)

References. Mathews and Walker Chap 6, Zauderer Chap 3.