The Method of Characteristics

Sean Kim

October 23, 2024

Introduction

In mathematics, a general principle that helps with tackling difficult problems is reducing said task into easier subtasks which can be handled individually rather than taking on the whole task as a whole. Indeed, this statement applies to the study of partial differential equations, as solving PDEs directly is often very difficult. The technique in question is known as the ***method of characteristics***, which reduces a given PDE to a system of ODEs, which are in principle, much easier than the single PDE.

The general principle behind the method of characteristics is based on the multivariate chain rule, which states that for a differentiable function $u : \mathbb{R}^n \to \mathbb{R}$:

$$\frac{\mathrm{d}u}{\mathrm{d}s} = \dot{u} = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}s} = \sum_{i=1}^{n} u_{x_i} \dot{x}_i,$$

for some parameterization by $s \in \mathbb{R}$. We use the dot notation above to mean the same thing as derivative of a real function with respect to a parameter (s in this case).

The general principle behind the method of characteristics is based on the multivariate chain rule, which states that for a differentiable function $u : \mathbb{R}^n \to \mathbb{R}$:

$$\frac{\mathrm{d}u}{\mathrm{d}s} = \dot{u} = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}s} = \sum_{i=1}^{n} u_{x_i} \dot{x}_i,$$

for some parameterization by $s \in \mathbb{R}$. We use the dot notation above to mean the same thing as derivative of a real function with respect to a parameter (s in this case).

An Easy Example

Suppose we want to solve the following PDE:

$$\begin{cases} -yu_x + xu_y = u & (x, y) \in U, \\ u = g & (x, y) \in \Gamma \subset \partial U, \end{cases}$$

where $U = \{x, y > 0\}$ and $\Gamma = \{x > 0, y = 0\}$.

An Easy Example (cont.)

Suppose that we have some parameterization of x and y by s. First, notice that by the chain rule we have

$$\dot{u} = u_x \dot{x} + u_y \dot{y}.$$

Furthermore, due to the linear structure of our given PDE, this suggests that we want to set $\dot{x} = -y$ and $\dot{y} = x$, which would give us

$$\dot{u} = -y u_x + x u_y = u.$$

An Easy Example (cont.)

Suppose that we have some parameterization of x and y by s. First, notice that by the chain rule we have

$$\dot{u} = u_x \dot{x} + u_y \dot{y}.$$

Furthermore, due to the linear structure of our given PDE, this suggests that we want to set $\dot{x} = -y$ and $\dot{y} = x$, which would give us

$$\dot{u} = -y u_x + x u_y = u.$$

Hence, we know that if u satisfies the following system, it can generate solutions to the desired PDE:

$$\begin{cases} \dot{x} & = -y, \quad \dot{y} = x, \\ \dot{u} & = u. \end{cases}$$

This system of ODEs are called the ***characteristic equations*** of the PDE.

Geometric Interpretation of the Method of Characteristics

We demonstrate the case of two variables for simplicity, though this geometric reasoning may be easily generalized to n dimensions. Consider the general first order quasilinear PDE described by:

$$a(x, y)u_x + b(x, y)u_y = c(x, y, u).$$

Suppose that $z(x, y) = u$ is a solution to the PDE. Consider the graph of z in $\mathbb{R}^3$. We know that the normal vectors to surface are of the form $\langle z_x, z_y, -1\rangle$.

Geometric Interpretation of the Method of Characteristics

We demonstrate the case of two variables for simplicity, though this geometric reasoning may be easily generalized to n dimensions. Consider the general first order quasilinear PDE described by:

$$a(x,y)u_x + b(x,y)u_y = c(x,y,u).$$

Suppose that $z(x,y) = u$ is a solution to the PDE. Consider the graph of z in $\mathbb{R}^3$. We know that the normal vectors to surface are of the form $\langle z_x, z_y, -1\rangle$. However, for z to be a solution to the PDE, we require that

$$\langle a(x,y), b(x,y), c(x,y,z)\rangle \cdot \langle z_x, z_y, -1\rangle = 0,$$

which states that $\langle a(x,y), b(x,y), c(x,y,z)$ must lie in the tangent planes to the graph of z. This means that the graph of z lies in the union of the integral curves to the vector field, or the spaces of solutions of the natural ODEs that arise.

Geometric Interpretation of the Method of Characteristics

To elaborate on the statement above, we require these characteristic curves to intersect the boundary conditions. That is, if we have a PDE of one spatial variable x and one temporal variable t such that

$$u = g \text{ if } (x, t) \in \Gamma \subset \partial U$$

for some given g on the boundary (which would be $\{t = 0\}$ if our domain is $\mathbb{R} \times [0, \infty)$), we require the characteristic curves to intersect $(x, t, u) = (x^0, 0, g)$ in $\mathbb{R}^3$ for each x^0. Hence, as we draw out each characteristic curve varying the x^0 value on the boundary, we propagate lines across space along which a solution u of the PDE has some degree of constancy. Often, we project the characteristic curves onto the xt-plane for simplicity, which we will visualize later.

The Transport Equation

Let us illustrate the principles of the method of characteristics by solving the transport equation:

$$\begin{cases} u_t + b \cdot Du = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

The Transport Equation

Let us illustrate the principles of the method of characteristics by solving the transport equation:

$$\begin{cases} u_t + b \cdot Du = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

Clearly, the characteristic equations are

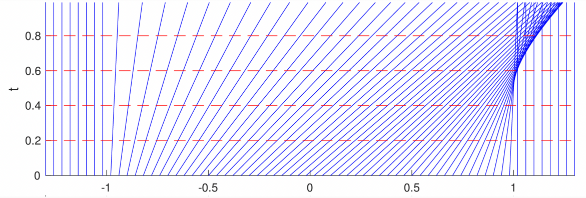
$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1, \ \frac{\mathrm{d}x}{\mathrm{d}s} = b, \ \frac{\mathrm{d}u}{\mathrm{d}s} = 0,$$

which implies that we may identify our parameter s with the time t. Furthermore, our characteristic curves are lines of the form $x = bt + x^0$ for some initial x^0 on the boundary. Since u is constant on the characteristic curves,

$$u(x^0, 0) = g(x^0) \implies u(x - bt, 0) = g(x - bt) \implies u(x, t) = g(x - bt).$$

In the transport equation, we saw that the method of characteristics successfully generated the general solution to the IVP, and we know that this solution exists and is unique because the characteristic curves span the whole space. However, we must be careful to ensure that these characteristic lines do not intersect. In the transport equation, the caveat was that the slope every characteristic curve was the same, meaning that all the lines would be parallel, determined by our coefficient $b \in \mathbb{R}^n$. This is not necessarily true in general, and there are situations where the lines may cross, forming ***shock waves*** where they intersect, where we may lose smoothness, continuity, and uniqueness of solutions.

In the transport equation, we saw that the method of characteristics successfully generated the general solution to the IVP, and we know that this solution exists and is unique because the characteristic curves span the whole space. However, we must be careful to ensure that these characteristic lines do not intersect. In the transport equation, the caveat was that the slope every characteristic curve was the same, meaning that all the lines would be parallel, determined by our coefficient $b \in \mathbb{R}^n$. This is not necessarily true in general, and there are situations where the lines may cross, forming ***shock waves*** where they intersect, where we may lose smoothness, continuity, and uniqueness of solutions.



Now that we are more familiar with the power and limitations of the method of characteristics, we seek a general formulation of the technique for nonlinear first-order PDEs.

Now that we are more familiar with the power and limitations of the method of characteristics, we seek a general formulation of the technique for nonlinear first-order PDEs.

The Method of Characteristics

Consider the following problem:

$$\begin{cases} F(Du, u, x) = 0 & \text{in } U \subset \mathbb{R}^n, \\ u = g & \text{on } \Gamma \subset \partial U. \end{cases}$$

Following Evans's derivation of the method, we parameterize x with s, such that $x(s) = (x_1(s), \ldots, x_n(s))$. Supposing that u is a C^2 solution to the PDE, we define

$$z(s) = u(x(s)), \quad p(s) = Du(x(s)),$$

where $p^i = u_{x_i}(x)$.

The Method of Characteristics

By differentiating the original PDE and calculating the derivatives of p, we may reduce any second order derivative terms in $\dot{p}$, and we obtain the following $2n+1$ characteristic equations for our PDE:

$$\begin{cases} \dot{p}(s) &= -D_x F - D_z F p(s), \\ \dot{z}(s) &= D_p F \cdot p(s), \\ \dot{x}(s) &= D_p F. \end{cases}$$

where F is in terms of $p(s), z(s)$, and $x(s)$. Furthermore, we know that

$$F(p(s), z(s), x(s)) \equiv 0$$

by the statement of the PDE. We call p and x the characteristics of the PDE, and clearly, by integrating appropriately, we may always derive an implicit solution for u.

Characteristics for Conservation Laws

To demonstrate shock wave phenomena and the limitations of the method of characteristics, we now turn to the scalar conservation laws:

$$\begin{cases} u_t + \operatorname{div} F(u) = u_t + F'(u) \cdot Du = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases}$$

F is a vector field from $\mathbb{R}$ to $\mathbb{R}^n$, and we set $t = x_{n+1}$. To account for time, we write $q = (p, p_{n+1})$ and $y = (x, t)$. Hence, rewriting the PDE in the appropriate notation,

$$G(q, z, y) = p_{n+1} + F'(z) \cdot p.$$

Hence, we calculate the characteristic equations of the PDE to be

$$\dot{x}(s) = F'(z(s)),\ \dot{x^{n+1}}(s) = 1,\ \dot{z}(s) = 0.$$

Characteristics for Conservation Laws

By the boundary condition and the fact that $\dot{z}(s) = 0$ on the characteristic curves, we know that

$$z(s) = g(x^0)$$

for some x^0 on the boundary. Furthermore, the ODEs for x imply that

$$x(s) = F'(g(x^0))s + x^0, \text{ and } y(s) = (x(s), s) = (F'(g(x^0))s + x^0, s),$$

so the trajectory of each of the characteristic curves (alone which u is constant) represents a straight line in the xt-plane.

Characteristics for Conservation Laws

By the boundary condition and the fact that $\dot{z}(s) = 0$ on the characteristic curves, we know that

$$z(s) = g(x^0)$$

for some x^0 on the boundary. Furthermore, the ODEs for x imply that

$$x(s) = F'(g(x^0))s + x^0, \text{ and } y(s) = (x(s), s) = (F'(g(x^0))s + x^0, s),$$

so the trajectory of each of the characteristic curves (alone which u is constant) represents a straight line in the xt-plane. We may also derive an implicit formula for u by noticing $s = t$ and hence

$$u(x^0, 0) = g(x^0) \implies u(x, t) = u(x - tF'(g(x^0)), 0) = g(x(t) - tF'(g(x^0))),$$

and $u = g(x - tF'(u))$ since u is constant along the characteristic lines.

Integral Solutions

Consider the scalar conservation laws with one spatial variable.

$$\begin{cases} u_t + F(u)_x = 0 & \text{in } \mathbb{R} \times (0, \infty) \\ u = g & \text{on } \mathbb{R} \times \{t = 0\} \end{cases}$$

We now attempt to analyze the issues that arise when two characteristic curves intersect and form a shock wave. First, We develop a notion of a weak solution of the PDE, and derive the Rankine-Hugoniot condition.

Integral Solutions

Consider the scalar conservation laws with one spatial variable.

$$\begin{cases} u_t + F(u)_x = 0 & \text{in } \mathbb{R} \times (0, \infty) \\ u = g & \text{on } \mathbb{R} \times \{t = 0\} \end{cases}$$

We now attempt to analyze the issues that arise when two characteristic curves intersect and form a shock wave. First, We develop a notion of a weak solution of the PDE, and derive the Rankine-Hugoniot condition.
Let $v : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ be in C_c^∞. We call v a test function, and we attempt to move the derivatives in the conservation law to v, so that we can relax the differentiability restrictions on u.

Integral Solutions

In light of the fundamental lemma of the calculus of variations, consider the following integral:

$$\int_0^\infty \int_{-\infty}^\infty (u_t + F(u)_x)v = 0.$$

By splitting the integral then integrating both integrals by parts and considering the domain on which v is supported, we obtain the identity

$$\int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \,\mathrm{d}x \,\mathrm{d}t + \int_{-\infty}^\infty gv \,\mathrm{d}x \,|_{t=0} = 0.$$

Integral Solutions

In light of the fundamental lemma of the calculus of variations, consider the following integral:

$$\int_0^\infty \int_{-\infty}^\infty (u_t + F(u)_x)v = 0.$$

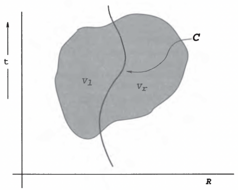
By splitting the integral then integrating both integrals by parts and considering the domain on which v is supported, we obtain the identity

$$\int_0^\infty \int_{-\infty}^\infty uv_t + F(u)v_x \,\mathrm{d}x\,\mathrm{d}t + \int_{-\infty}^\infty gv\,\mathrm{d}x\,|_{t=0} = 0.$$

Note that this allows us to generalize the domain to which u belongs; u now need not be differentiable, but rather, we only require u to be essentially bounded. Hence, we say that u is an integral solution to the conservation law if it is in $L^\infty(\mathbb{R} \times (0, \infty))$ and satisfies the aforementioned identity for all test functions.

Rankine-Hugoniot Condition

Consider the following diagram from Evans:



Rankine-Hugoniot condition

We see that V is a region in $\mathbb{R} \times (0, \infty)$ such that the smooth curve C divides it into two parts, V_l and V_r. We suppose that u is smooth on either side of V. By integrating the integral solution identity by parts with v compactly supported in V_l, we deduce

$$0 = -\int_0^\infty \int_{-\infty}^\infty [u_t + F(u)_x] v \, \mathrm{d}x \, \mathrm{d}t \, .$$

Rankine-Hugoniot Condition

As a result, we find that

$$u_t + F(u)_x = 0$$

in V_l, and similar reasoning extends this to V_r. However, if v is compactly supported in V but does not vanish along the curve C, the case is more subtle. We find that

$$0 = \iint_{V_l} uv_t + F(u)v_x \,\mathrm{d}x\,\mathrm{d}t + \iint_{V_r} uv_t + F(u)v_x \,\mathrm{d}x\,\mathrm{d}t\,,$$

and calculating the first integral by parts gives

$$-\iint_{V_l} [u_t + F(u)_x] v \,\mathrm{d}x\,\mathrm{d}t + \int_C (u_l \nu^2 + F(u_l)\nu^1) v \,\mathrm{d}l = \int_C (u_l \nu^2 + F(u_l)\nu^1) v \,\mathrm{d}l\,,$$

where $\nu = (\nu^1, \nu^2)$ is the normal vector on C facing into V_r, and the subscript l denotes the limit from the left (V_l).

Rankine-Hugoniot Condition

Using similar reasoning for V_r, we find that

$$\iint_{V_r} uv_t + F(u)v_x \, \mathrm{d}x \, \mathrm{d}t = -\int_C (u_r\nu^2 + F(u_r)\nu^1)v \, \mathrm{d}l \,,$$

and adding the two integrals gives that

$$0 = \int_C [(F(u_l) - F(u_r))\nu^1 + (u_l - u_r)\nu^2]v \, \mathrm{d}l \,.$$

By the fundamental lemma of the calculus of variations, we have

$$[\![F(u)]\!]\nu^1 + [\![u]\!]\nu^2 = 0$$

along C, where $[\![F(u)]\!]$ is the jump in $F(u)$ from the left to right of C and likewise for $[\![u]\!]$.

Rankine-Hugoniot Condition

Explicitly parameterizing the curve C with some smooth function s and calculating the outward facing normal's components, we obtain the equality

$$[\![F(u)]\!] = \sigma[\![u]\!],$$

where $\sigma = \dot{s}$, or the speed of the curve C. This is known as the Rankine Hugoniot condition along the shock curve C. It tells us information about how the value of the functions jump across a discontinuity, and can correspond physically to the conservation of momentum, energy, or mass before and after a shock wave.