Euler equation: Multivariable problem

April 1, 2019

Contents

1 Reminder of multivariable calculus

This part deals with multivariable variational problems that describe minimal surface areas, equilibria and dynamics of continua, optimization of shapes, etc. These problems require minimization of an integral of a multivariable function and its gradient over a region Ω of a plane or space. The arguments of the Lagrangian are vector u of minimizers and matrix $Du = \nabla u$ of gradients of these minimizers. Analysis of these problems requires differentiation with respect to vectors and matrices. First, we recall several formulas of vector (multivariable) calculus which will be commonly used in the following chapters.

1.1 Vector differentiation

We remind the definition of a vector derivative or derivative of a scalar function with respect to a vector argument $a \in R^n$.

Definition 1.1 If $\phi(a)$ is a scalar function of a column vector argument $a = (a_1, \ldots, a_n)^T$, then the derivative $\frac{d\phi}{da}$ is a row vector

$$\frac{d\phi}{da} = \left(\frac{d\phi}{da_1}, \ldots, \frac{d\phi}{da_n}\right) \quad \text{if} \quad a = \begin{pmatrix} a_1 \\ \ldots \\ a_n \end{pmatrix} \tag{1}$$

assuming that all partial derivatives exist.

This definition comes from consideration of differential $d\phi$ of the scalar function $\phi(a)$:

$$d\phi(a) = \phi(a + da) - \phi(a) = \frac{d\phi(a)}{da} \cdot da + o(\|a\|)$$

and the left-hand side is a scalar. Here, da is a column vector, therefore the first multiplier $\frac{d\phi}{da}$ is a row vector defined in (1)

Examples of vector differentiation The next examples show the calculation of derivative for several often used functions. The results can be checked by straightforward calculations. We assume here that $a \in R^n$.

1. If $\phi(a) = |a|^2 = a_1^2 + \ldots a_n^2$, the derivative is

$$\frac{d}{da}|a|^2 = 2a^T$$

2. The vector derivative of the euclidean norm $|a|$ of vector a is a row vector b,

$$b = \frac{d}{da}\sqrt{|a|^2} = \frac{a^T}{\sqrt{|a|^2}} = \frac{a^T}{|a|}$$

Observe that b is codirected with a and has unit length.

3. The derivative of a scalar product $c \cdot a$, where c is an n-dimensional vector, $c \in R^n$, is equal to c:
$$\frac{d}{da} c^T a = c^T$$
Similarly, if C is a $k \times n$ matrix, derivative of a product Ca equals C,
$$\frac{d}{da} Ca = C$$

4. Derivative of a quadratic form $a^T Ca$ where C is a symmetric matrix, equals
$$\frac{d}{da} a^T Ca = 2a^T C = 2(Ca)^T.$$

Directional derivative Let ϕ_ν be a directional derivative of a scalar function ϕ in a direction ν: $\phi_\nu = \nu \cdot \nabla\phi$. Partial derivative of $F(\nabla\phi)$ with respect to ϕ_ν is defined as:
$$\frac{\partial F}{\partial \phi_\nu} = \frac{\partial F}{\partial \nabla \phi} \cdot \nu \tag{2}$$

Gradient of a scalar function. Integrability conditions If $u = u(x_1, \dots x_d)$, $x \in R^d$, is a function of position vector x then the gradient of u is defined as a vector denoted ∇u or Du, $\nabla u = \frac{\partial u}{\partial x_i}$, or, in elements,
$$\nabla u = Du = \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \dots \\ \frac{\partial u}{\partial x_d} \end{pmatrix} \tag{3}$$

The d entries of the vector $v = \nabla u$ depend on one scalar function u and therefore are connected. Indeed
$$\frac{\partial v_i}{\partial x_j} = \frac{\partial^2 u}{\partial x_i \partial x_j}, \quad \frac{\partial v_j}{\partial x_i} = \frac{\partial^2 u}{\partial x_j \partial x_i},$$
Because the mixed partials are equal, we find integrability conditions
$$\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}, \quad i, j = 1, \dots, d \tag{4}$$
They are conveniently expressed through the curl-operator $(\nabla \times)$ which acts on a vector W as follows:
$$\nabla \times W = \begin{pmatrix} 0 & -\frac{\partial}{\partial x_3} & \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} & 0 & -\frac{\partial}{\partial x_1} \\ -\frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} & 0 \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} = \begin{pmatrix} -\frac{\partial W_2}{\partial x_3} + \frac{\partial W_3}{\partial x_2} \\ \frac{\partial W_1}{\partial x_3} - \frac{\partial W_3}{\partial x_1} \\ -\frac{\partial W_1}{\partial x_2} + \frac{\partial W_2}{\partial x_1} \end{pmatrix}$$
The compatibility conditions (4) can be written as
$$\nabla \times \nabla u = 0 \tag{5}$$
for any $u(x)$.

Gradient of a vector function If $u = (u_1, \dots, u_n)$ is a vector function, $u_j = u_j(x_1, \dots x_d)$, $x \in R^d$, then the gradient of u is defined as a $d \times n$ matrix denoted ∇u or Du, $\nabla u = \frac{\partial u_j}{\partial x_i}$, or, in elements,

$$\nabla u = Du = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_2}{\partial x_1} & \dots & \frac{\partial u_n}{\partial x_1} \\ \dots & \dots & \dots & \dots \\ \frac{\partial u_1}{\partial x_d} & \frac{\partial u_2}{\partial x_d} & \dots & \frac{\partial u_n}{\partial x_d} \end{pmatrix} = (\nabla u_1 | \nabla u_2 | ... | \nabla u_n) \tag{6}$$

The columns of this matrix are gradients of the components of the vector function u.

1.2 Matrix differentiation

Similarly to the vector differentiation we define matrix differentiation considering a scalar function $\phi(A)$ of a matrix argument A. As in the vector case, the definition is based on the notion of scalar product.

Definition 1.2 The scalar product a.k.a. the convolution of the $n \times m$ matrix A and $m \times n$ matrix B is defined as following

$$A : B = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} b_{ji}.$$

One can check the formula

$$A : B = \mathrm{Tr}\,(A\,B) \tag{7}$$

that brings the convolution into the family of familiar matrix operations.

The convolution allows us to for calculate the increment of a *matrix-differentiable* function of a matrix argument caused by variation of this argument:

$$d\phi(A) = \phi(A + dA) - \phi(A) = \frac{d\phi(A)}{dA} : dA + o(\|dA\|).$$

and to give the definition of the matrix-derivative.

Definition 1.3 The derivative of a scalar function ϕ by an $n \times m$ matrix argument A is an $m \times n$ matrix $D = \frac{d\phi}{dA}$ with elements

$$D_{ij} = \frac{\partial \phi}{\partial a_{ji}}$$

where a_{ij} is the ij-element of A.

In element form, the definition becomes

$$\frac{d\phi}{dA} = \begin{pmatrix} \frac{\partial \phi}{\partial a_{11}} & \frac{\partial \phi}{\partial a_{21}} & \dots & \frac{\partial \phi}{\partial a_{m1}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \phi}{\partial a_{1n}} & \frac{\partial \phi}{\partial a_{2n}} & \dots & \frac{\partial \phi}{\partial a_{mn}} \end{pmatrix} \tag{8}$$

Examples of matrix differentiation Next examples show the derivatives of several often used functions of matrix argument.

1. As the first example, consider $\phi(A) = \operatorname{Tr} A = \sum_{i=1}^{n} a_{ii}$. Obviously,

$$\frac{d\phi}{da_{ij}} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases},$$

therefore the derivative of the trace is the unit matrix,

$$\frac{d}{dA} \operatorname{Tr} A = I.$$

2. Using definition of the derivative, we easily compute the derivative of the scalar product or convolution of two matrices,

$$\frac{d(A:B)}{dA} = \operatorname{Tr} A\ B = B.$$

3. Assume that A is a square $n \times n$ matrix. The derivative of the quadratic form $x^T A x = \sum_{i,j=1}^{n} x_i x_j a_{ij}$ with respect to matrix A is an $n \times n$ dyad matrix

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

4. Compute the derivative of the determinant of matrix A for $A \in R^{n \times n}$. Notice that the determinant linearly depends on each matrix element,

$$\det A = a_{ij} M_{ij} + \text{ constant}(a_{ij})$$

where M_{ij} is the minor of the matrix A obtained by eliminating the ith row and the jth column; it is independent of a_{ij}. Therefore,

$$\frac{\partial \det A}{\partial a_{ij}} = M_{ij}$$

and the derivative of $\det A$ is the matrix M of minors of A,

$$\frac{d}{dA} \det A = M = \begin{pmatrix} M_{11} & \dots & M_{1n} \\ \dots & \dots & \dots \\ M_{n1} & \dots & M_{nn} \end{pmatrix}$$

Recall that the inverse matrix A^{-1} can be conveniently expressed through these minors as $A^{-1} = \frac{1}{\det A} M$, and rewrite the result as

$$\frac{d}{dA} \det A = (\det A) A^{-1}$$

We can rewrite the result once more using the logarithmic derivative $\frac{d}{dx} \log f(x) = \frac{f'(x)}{f(x)}$. The derivative becomes more symmetric,

$$\frac{d}{dA} (\log \det A) = A^{-1}.$$

Remark 1.1 If A is symmetric and positively defined, we can bring the result to a perfectly symmetric form

$$\frac{d}{d\log A}(\log\det A) = I$$

Indeed, we introduce the matrix logarithmic derivative similarly to the logarithmic derivative of a real positive argument,

$$\frac{df}{d\log x} = x\frac{df}{dx},$$

which reads

$$\frac{df}{d\log A} = A\frac{df}{dA}.$$

Here, $\log A$ is the matrix that has the same eigenvectors as A and the eigenvalues equal to logarithms of the corresponding eigenvalues of A. Notice that $\log\det A$ is the sum of logarithms of the eigenvalues of A,

$$\log\det A = \text{ Tr } \log A$$

. Notice also that when the matrix A is symmetric and positively defined which means that the eigenvalues of A are real and positive, the logarithms of the eigenvalues are real.

5. Using the chain rule, we compute the derivative of the trace of the inverse matrix:

$$\frac{d}{dA} \text{ Tr } A^{-1} = -A^{-2}.$$

6. Similarly, we compute the derivative of the quadratic form associated with the inverse matrix:

$$\frac{d}{dA}x^T A^{-1}x = -xA^{-2}x^T.$$

Remark 1.2 (About the notations) The historical Leibnitz notation $g = \frac{\partial f}{\partial z}$ for partial derivative is not the most convenient one and can even be ambiguous. Indeed, the often used in one-variable variational problems partial $\frac{\partial f}{\partial u'}$ becomes in multivariable problem the partial of the partials $\frac{\partial u}{\partial x}$. Since there is no conventional analog for the symbol $'$ in partial derivatives, we need a convenient way to express the fact that the argument z of differentiation can itself be a partial derivative like $z = \frac{\partial u_1}{\partial x_2}$. If we were substitute this expression for z into $\frac{\partial f}{\partial z}$, we would arrive at an a bit awkward expression

$$g = \frac{\partial f}{\partial \frac{\partial u_1}{\partial x_2}}$$

(still used in Gelfand & Fomin) which replaces the expression $\frac{\partial f}{\partial u'}$ used in one-variable variational problem.

There are several ways to fix the inconvenience. To keep analogy with the one-variable case, we use the vector of partials $\frac{\partial f}{\partial(\nabla u)}$ in the place of $\frac{\partial F}{\partial u'}$. If needed, we specify a component of this vector, as follows

$$g = \left[\frac{\partial f}{\partial(\nabla u_1)}\right]_2$$

Alternatively, we could rename the partial derivatives of u with a single indexed array D_{ij} arriving at the formula of the type

$$g = \frac{\partial f}{\partial D_{12}}, \quad \text{where } D_{12} = \frac{\partial u_1}{\partial x_2}.$$

or use comma to show the derivative

$$g = \frac{\partial f}{\partial u_{1,2}}, \quad \text{where } u_{1,2} = \frac{\partial u_1}{\partial x_2}.$$

The most radical and logical solution (which we do not dare to develop in the textbook) replaces Leibnitz notation with something more convenient, namely with Maple-like notation

$$g = D(f, D(u_1, x_2))$$

Remark 1.3 (Ambiguity in notations) A more serious issue is the possible ambiguity of partial derivative with respect to one of independent coordinates. The partial $\frac{\partial}{\partial x}$ means the derivative upon the explicitly given argument x of a function of the type $F(x, u)$. If the argument x is one of the independent coordinates, and if u is a function of these coordinates, in particular of x (as it is common in calculus of variations problems), the same partial could mean $\frac{\partial F}{\partial x} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x}$. To fix this, we need to specify whether we consider u as a function of x $u = u(x)$ or as an independent argument, which could make the notations awkward.

For this reason, we always assign the symbol x for a vector of independent variables (coordinates). When differentiation with respect to independent coordinates is considered, we use the gradient notations as ∇u. Namely, the vector is introduced

$$\nabla F(u, x) = \begin{pmatrix} \frac{\partial F}{\partial x_1} \\ .. \\ \frac{\partial F}{\partial x_d} \end{pmatrix} + \begin{pmatrix} \frac{\partial F}{\partial u}\frac{\partial u}{\partial x_1} \\ .. \\ \frac{\partial F}{\partial u}\frac{\partial u}{\partial x_d} \end{pmatrix}$$

where $\frac{\partial}{\partial x_k}$ always means the derivative upon explicit variable x. The partials corresponds to components of this vector. If necessary, we specify the argument of the gradient, as follows ∇_ξ.

1.3 Multidimensional integration

Change of variables Consider the integral

$$I = \int_\Omega f(x)\, dx$$

and assume that $x = x(\xi)$, or in coordinates

$$x_i = x_i(\xi_1, \ldots, \xi_d), \quad i = 1, \ldots, d$$

In the new coordinates, the domain Ω is mapped into the domain Ω_ξ and the volume element dx becomes $\det J \, d\xi$ where J is the Jacobian, the matrix gradient

$$J = \nabla_\xi x = \{J_{ij}\}, \quad J_{ij} = \frac{\partial x_i}{\partial \xi_j}, \quad i, j = 1, \ldots, d.$$

The integral I becomes

$$I = \int_{\Omega_\xi} f(x(\xi))(\det \nabla_\xi x) \, dx \tag{9}$$

The change of variables in the multivariable integrals is analogous to the one-dimensional case.

Green's formula The Green's formula is a multivariable analog of the Leibnitz formula a.k.a. the fundamental theorem of calculus. For a differentiable in the domain Ω vector-function $a(x)$ it has the form

$$\int_\Omega \nabla \cdot a \, dx = \int_{\partial\Omega} a \cdot \nu \, ds \tag{10}$$

Here, ν is the outer normal to Ω.

Integration by parts We will use multivariable analogs of the integration by parts. Suppose that $b(x)$ is a differentiable scalar function in Ω and $a(x)$ is a differentiable vector field in Ω. Then the following generalization of integration by parts holds

$$\int_\Omega (a \cdot \nabla b) \, dx = -\int_\Omega (b \nabla \cdot a) \, dx + \int_{\partial\Omega} (a \cdot \nu) b \, ds \tag{11}$$

The formula follows from the differential identity (differentiation of a product)

$$a \cdot \nabla b + b \nabla \cdot a = \nabla \cdot (ba)$$

and Green's formula

$$\int_\Omega (\nabla \cdot c) dx = \int_{\partial\Omega} (c \cdot \nu) \, ds$$

A similar formula holds for curls of two differentiable in Ω vector fields a and c:

$$\int_\Omega (c \cdot \nabla \times b) \, dx = \int_\Omega (b \cdot \nabla \times c) \, dx - \int_{\partial\Omega} (c \times b \cdot \nu) \, ds \tag{12}$$

It immediately follows from the Green's formula and the identity

$$\nabla \cdot (c \times b) = b \cdot \nabla \times c - c \cdot \nabla \times b$$

2 Euler equations for multiple integrals

2.1 Euler equation

Consider the simplest problem of multivariable calculus of variation: Minimize an integral of a twice differentiable Lagrangian $F(x, u, \nabla u)$ over a regular bounded domain Ω with a smooth boundary $\partial\Omega$. The Lagrangian F depends on the minimizer u and its gradient ∇u with the function u taking prescribed values u_0 on the boundary $\partial\Omega$,

$$\min_{u:\, u|_{\partial\Omega}=u_0} I(u), \quad I(u) = \int_\Omega F(x, u, \nabla u) dx \tag{13}$$

As in the one-variable version (see Chapter 2), the Euler equation expresses the stationarity of the functional I with respect to the variation of u. To derive the Euler equation, we consider the variation δu of the minimizer u and the increment of the functional $\delta I = I(u + \delta u) - I(u)$. We assume that for any x, the variation δu is localized in an ϵ-neighborhood of point x, twice differentiable, and small: the norm of δu of its gradient goes to zero if $\epsilon \to 0$,

$$\delta u(x + z) = 0,\ \forall z:\ |z| \geq \epsilon, \quad |\nabla(\delta u)| < C\epsilon,\ \forall x \tag{14}$$

If u is a minimizer, the increment δI must be nonnegative, $\delta I(u, \delta u) \geq 0$, $\forall \delta u$.

Increment When the variation δu and its gradient are both infinitesimal and F is twice differentiable, we can linearize the perturbed Lagrangian:

$$\begin{aligned} F(x, u + \delta u, \nabla(u + \delta u)) \ &= \ F(x, u, \nabla u) + \frac{\partial F(x, u, \nabla u)}{\partial u}\delta u \\ &+ \ \frac{\partial F(x, u, \nabla u)}{\partial \nabla u}\delta\nabla u + o(||\delta u||, ||\nabla(\delta u)||) \end{aligned}$$

Here, the term $\frac{\partial F(x,u,\nabla u)}{\partial \nabla u}$ denotes the vector of the partial derivatives of F with respect to partial derivatives of u,

$$\frac{\partial F(x, u, \nabla u)}{\partial \nabla u} = \left[\frac{\partial F(x, u, \nabla u)}{\partial\left(\frac{\partial u}{\partial x_1}\right)}, \ldots, \frac{\partial F(x, u, \nabla u)}{\partial\left(\frac{\partial u}{\partial x_n}\right)}\right].$$

Substitution of the linearized Lagrangian into the expression for δI results in the following expression:

$$\delta I = \int_\Omega \left(\frac{\partial F}{\partial u}\delta u + \underline{\frac{\partial F}{\partial \nabla u} \cdot \delta\nabla u}\right) dx + o(||\delta u||, ||\nabla\delta u||).$$

Next, we transform the underlined term. Interchanging two linear operators of variation and differentiation, $\delta\nabla u = \nabla\delta u$, and performing integration by parts (see (11)), we obtain

$$\int_\Omega \left(\frac{\partial F}{\partial \nabla u} \cdot \nabla(\delta u)\right) dx = -\int_\Omega \delta u \left(\nabla \cdot \frac{\partial F}{\partial \nabla u}\right) dx + \int_{\partial\Omega} \delta u \left(\frac{\partial F}{\partial \nabla u} \cdot n\right) ds$$

so that

$$\delta I = \int_\Omega \left(\frac{\partial F}{\partial u} - \nabla \cdot \frac{\partial F}{\partial \nabla u} \right) \delta u \, dx + \int_{\partial\Omega} \delta u \left(\frac{\partial F}{\partial \nabla u} \cdot n \right) \, ds$$

The coefficient by δu in the first integral is called the *variational derivative in* Ω or the *sensitivity function*:

$$S_F(u) = \frac{\partial F}{\partial u} - \nabla \cdot \left(\frac{\partial F}{\partial \nabla u} \right) \tag{15}$$

The coefficient by δu in the boundary integral is called the *variational derivative on the boundary* $\partial\Omega$:

$$S_F^\partial(u, n) = \frac{\partial F}{\partial \nabla u} \cdot n = \frac{\partial F}{\partial \left(\frac{\partial u}{\partial n} \right)} \tag{16}$$

Hence, we represent the linearized increment δI as a sum of two terms:

$$\delta I = \int_\Omega S_F(u) \delta u \, dx + \int_{\partial\Omega} S_F^\partial(u, n) \delta u \, ds. \tag{17}$$

Notice that we use (2) to write the last term in (16).

Stationarity The condition $\delta I \geq 0$ and the arbitrariness of variation δu in the domain Ω and possibly on its boundary $\partial\Omega$ leads to the stationarity condition in a form of differential equation:

$$S_F(u) = 0 \quad \text{or} \quad -\nabla \cdot \frac{\partial F}{\partial \nabla u} + \frac{\partial F}{\partial u} = 0 \quad \text{in } \Omega \tag{18}$$

with the boundary condition

$$S_F^\partial(u, n) \, \delta u = 0 \quad \text{on } \partial\Omega \tag{19}$$

Equation (18) with the boundary condition (19) is the **Euler-Lagrange equation** for variational problems dealing with multiple integrals. Notice that we keep δu in the expression for the boundary condition. This allows us to either assign u on the boundary or leave it free, which corresponds to two different types of boundary condition.

The main boundary condition In the considered simplest problem, the partial differential equation (18) is given in Ω with the boundary condition $u = u_0$. The boundary term (19) of the increment vanishes because the value of u on the boundary is prescribed, hence the variation δu is zero. This condition is called the *main boundary condition*. It is assigned independently of any variational requirements. When u is prescribed on some component of the boundary, we say that the main boundary condition is given; in this case the variation of u on this part of the boundary is zero, $\delta u = 0$, and (19) is satisfied.

Natural boundary condition If the value of u on the boundary is not specified, the term (19) supplies additional boundary condition. If no condition is prescribed on a boundary component, δu is an arbitrary function, and the natural condition

$$S_F^{\partial}(u,n) = \frac{\partial F}{\partial(\nabla u)} \cdot n = 0 \quad \text{or} \quad \frac{\partial F}{\partial u_n} = 0 \tag{20}$$

must be satisfied. The natural boundary condition follows from the minimization requirement; it must be satisfied to minimize the functional in (13).

Thus the boundary value problem in the domain Ω has one condition, main or natural, on each component of the boundary.

Remark 2.1 Notice that the stationarity and the natural boundary conditions are in direct analogy with corresponding conditions for a one-variable Euler equation. The derivative $\frac{d}{dx}$ with respect to the independent variable is replaced by ∇ or by $\nabla\cdot$. At the boundary, the derivative $\frac{du}{dx}$ is replaced by $\frac{\partial u}{\partial n} = \frac{\partial}{\partial \nabla u} \cdot n$. In the last case, the derivative with respect to x is changed to the directional derivative along the normal to the boundary.

2.2 Examples of Euler-Lagrange equations

Here, we give several examples of Lagrangians, the corresponding Euler equations, and natural boundary conditions

Example 2.1 (Laplace's equation) Consider a Lagrangian that quadratically depends on ∇u:

$$F = \frac{1}{2}\nabla u \cdot \nabla u$$

This Lagrangian corresponds to the energy of a linear conducting medium of unit conductivity. We compute the variational derivative of F:

$$\frac{\partial F}{\partial \nabla u} = \nabla u, \quad S_F(u) = -\nabla \cdot \frac{\partial F}{\partial \nabla u} = -\nabla \cdot \nabla u.$$

The stationarity condition or the Euler equation, $S_F(u) = 0$, is Laplace's equation:

$$-\nabla \cdot \nabla u = -\Delta u = 0$$

where (-Δ) is the Laplace operator, or the Laplacian. In the coordinate notation, Laplace's equation has the form:

$$S_F(u) = -\sum_{i=1}^{d} \frac{\partial^2 u}{\partial x_i^2} = 0.$$

The natural boundary condition is

$$S_F^{\partial}(u,n) = \nabla u \cdot n = \frac{\partial u}{\partial n} = 0$$

Notice, that if no values of function u is prescribed on the boundary, so that the natural boundary condition is posed on whole boundary $\partial\Omega$, then identically zero solution, $u = 0$, is the only solution of this problem. To get a non-trivial solution, the main boundary condition should be given on a part of the boundary or the Lagrangian should be modified.

Example 2.2 (Linear elliptic equation) We consider a more general Lagrangian corresponding to the energy density of a linear conducting heterogeneous anisotropic material:

$$F = \frac{1}{2}\nabla u \cdot A(x)\nabla u$$

Here $A(x) = \{A_{ij}(x)\}$ is a symmetric positive definite conductivity tensor that represents the material properties, and u is the potential such as temperature, electric potential, or concentration of particles. The steady state distribution of the potential minimizes the total energy or solves the variational problem (13) with the Lagrangian F. We comment on the derivation of this energy below in Section **??**. Here, we are concerned with the form of the stationarity condition for this Lagrangian. The variational derivative is the following:

$$\frac{\partial F}{\partial \nabla u} = A\nabla u, \quad S_F(u) = -\nabla \cdot \frac{\partial F}{\partial \nabla u} = -\nabla \cdot A\nabla u.$$

The stationarity condition (Euler equation) is the second-order elliptic equation:

$$S_F(u) = -\nabla \cdot A(x)\nabla u = 0$$

which in the coordinate notation, has the form:

$$S_F(u) = -\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial}{\partial x_i} A_{ij} \frac{\partial u}{\partial x_j} = 0.$$

The natural boundary condition is $S_F^{\partial}(u, n) = A\nabla u \cdot n = 0$.

In the general case of anisotropic conductivity given by a tensor A, main boundary condition is called the *Dirichlet boundary condition*:

$$u = u_0 \quad \text{on} \quad \partial\Omega$$

The natural boundary condition corresponding to the *homogeneous Neumann condition* is:

$$A(x)\frac{\partial u}{\partial n} = 0 \quad \text{on} \quad \partial\Omega$$

Notice that in this case, the directional derivative, not the normal derivative, is zero on the boundary.

When A is proportional to the unit matrix I, $A = \kappa(x)I$, where $\kappa > 0$ is the scalar conductivity, the Lagrangian becomes

$$F = \frac{\kappa(x)}{2}\nabla u \cdot \nabla u$$

The corresponding Euler equation

$$\nabla \cdot \kappa(x) \nabla u = 0$$

describes conduction process in a inhomogeneous isotropic medium with a spatially varying scalar conductivity function $\kappa(x) > 0$. Using coordinate notation, the equation is written as:

$$\sum_{i=1}^{d} \frac{\partial}{\partial x_i} k(x) \frac{\partial u}{\partial x_i} = 0$$

The natural boundary condition is called the *homogeneous Neumann condition*:

$$k(x) \frac{\partial u}{\partial n} = 0 \quad \text{on} \quad \partial\Omega$$

it can be simplified to $\frac{\partial u}{\partial n} = 0$.

Example 2.3 (Poisson and Helmholtz equations) Let us demonstrate that classical linear elliptic equations of mathematical physics originate from a variational problem of minimization of a quadratic Lagrangian. The Lagrangian of a form:

$$F = \frac{1}{2} |\nabla u|^2 - \frac{1}{2} a u^2 - b u \tag{21}$$

corresponds to the Euler equation $S_F(u) = 0$:

$$\Delta u + a u + b = 0$$

which is called the inhomogeneous Helmholtz equation. The natural boundary condition $\frac{\partial u}{\partial n} = 0$ is independent of a and b. If $a = 0$, the inhomogeneous Helmholtz equation degenerates into *Poisson equation*. If $b = 0$, it becomes homogeneous *Helmholtz equation*, and if $a = b = 0$ it degenerates into Laplace equation.

Example 2.4 (Nonlinear elliptic equation) Assume that the Lagrangian depends only on magnitude of the gradient:

$$F = \phi(|\nabla u|) \tag{22}$$

where ϕ is a monotonically increasing convex function, $\phi'(z) > 0$, $\forall z \in [0, \infty)$. Such Lagrangians describe the steady state conductivity or diffusion process in an isotropic nonlinear medium; u is the potential or concentration of diffusing particles.

Let us assume that $|\nabla u|$ does not turn to zero. The Euler equation is computed as

$$\nabla \cdot (\kappa(|\nabla u|) \nabla u) = 0, \quad \kappa(z) = \frac{\phi'(z)}{|z|}$$

Since $\phi' > 0$, the equation is elliptic. It also can be rewritten as two first-order equations

$$\nabla \cdot j = 0, \qquad j = \phi'(|\nabla u|) \frac{\nabla u}{|\nabla u|}$$

where j is a divergence free vector of *current*. The first equation expresses the equilibrium of the current density. The second equation is called the *constitutive relation*. It demonstrates the property of the material and characterizes the dependence of the current on the *field* ∇u. The coefficient $\frac{\phi'(|\nabla u|)}{|\nabla u|}$ is the conductivity of a nonlinear material; it depends on the magnitude of the field.

The natural boundary condition is

$$\frac{\phi'(|\nabla u|)}{|\nabla u|}\frac{\partial u}{\partial n} = 0$$

Because $\phi' > 0$, it simplifies to $\frac{\partial u}{\partial n} = 0$ and again expresses the vanishing of the normal derivative of u on the boundary.

In the next examples, we specify the function ϕ and obtain the variational form of well-studied nonlinear equations.

Example 2.5 (Nonlinear elliptic equation) The previous problem simplifies if the Lagrangian depends on the squared magnitude of the gradient:

$$F = \psi(|\nabla u|^2) \tag{23}$$

we assume again here that ψ is a monotonically increasing convex function. Differentiating F, we have:

$$\frac{\partial F}{\partial \nabla u} = 2\psi'(|\nabla u|^2)\nabla u$$

So that the Euler equation is

$$\nabla \cdot \left(\psi'(|\nabla u|^2)\nabla u\right) = 0$$

The conductivity of the nonlinear medium in this case, is $\psi'(|\nabla u|^2)$; as in the previous case, it depends on the magnitude of the field. Special case, when $\psi(z) = z$ results in Laplace equation.

Example 2.6 (p-Laplacian) Consider the Lagrangian that corresponds to special nonlinearity $\phi(z) = \frac{1}{p}z^p$ in (22)

$$F = \frac{1}{p}|\nabla u|^p \tag{24}$$

The Euler equation is:

$$\nabla \cdot \left(|\nabla u|^{p-2}\nabla u\right) = 0$$

The equation is called p-Laplacian. It degenerates into Laplace equation when $p = 2$.

Example 2.7 ($p = 1$) Another interesting case $p = 1$. We consider the two-dimensional case, $d = 2$. The Lagrangian becomes the norm of the gradient,

$$F = |\nabla u| = \sqrt{\left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2} \tag{25}$$

The corresponding Euler equation is:

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) = 0 \quad \text{in } \Omega$$

In this case, the isotropic nonlinear conductivity function is $|\nabla u|^{-1}$. Here again, we assume that $|\nabla u| \neq 0$, otherwise $|\nabla u|$ can be approximated as $|\nabla u| = \sqrt{|\nabla u|^2 + \beta}$ for some small parameter β.

Similar to the general case, the Euler equation can be written as a system of two first-order partial differential equations

$$j = \frac{\nabla u}{|\nabla u|}, \quad \nabla \cdot j = 0, \quad |j| = 1,$$

Observe that in this case $|j| = 1$. Here, the current j is codirected with ∇u and has the unit magnitude. In 2D case, any unit vector admits the representation

$$j = (j_1, j_2), \quad j_1 = \cos\theta, \quad j_2 = \sin\theta$$

where $\theta(x)$ is an unknown scalar function, that is defined by the first-order equation $\nabla \cdot j = 0$ or

$$-\sin\theta \frac{\partial \theta}{\partial x_1} + \cos\theta \frac{\partial \theta}{\partial x_2} = 0$$

Potential u is found from another first-order equation that states that j is parallel to ∇u, or $j \times \nabla u$=0. In the coordinate form, the equation becomes

$$\frac{\partial u}{\partial x_1} j_2 - \frac{\partial u}{\partial x_2} j_1 = 0$$

Notice the Euler equation is split into two first order equations.

2.3 Smooth approximation and continuation

As a first application of the multivariable extremal theory, consider a problem of approximation of a given scalar function f of several variables by a function u with assumed smoothness. The problem of approximation of a bounded, integrable, but may be discontinuous function $f(x)$, with x being in some subdomain $D \subset R^3$, by a smooth function $u(x)$ results in the variational problem

$$\min_u \frac{1}{2} \int_{R^3} \left((u - f)^2 + \epsilon^2 |\nabla u|^2 \right) dx$$

where the term $\epsilon^2 |\nabla u|^2$ represents penalization. If $\epsilon \ll 1$, the first term of the integrand prevails, and u accurately approximates f. As the parameter ϵ grows, the approximation becomes less accurate but the function u becomes more smooth. When $\epsilon \gg 1$, the approximation u tends to a constant function equal to the mean value of f.

The Euler equation for the approximation u is the inhomogeneous Helmholtz equation:

$$\epsilon^2\nabla^2 u - u = -f \quad \text{in } D, \quad \lim_{|x|\to\infty} u(x) = 0$$

This inhomogeneous Helmholtz problem can be explicitly solved using Green's function representation. In 3D case, we have:

$$u(x) = \int_{R^3} f(y)K(x-y)dy$$

Here $K(x-y)$ is the Green's function which satisfies the equation:

$$(\epsilon^2\nabla^2 - 1)K(r) = -\delta(r), \quad \lim_{|r|\to\infty} u(r) = 0$$

The Green's function for this Helmholtz problem for the whole R^3 is

$$K(r) = \frac{1}{4\epsilon^2\pi|r|}\exp\left(-\frac{|r|}{\epsilon}\right), \quad |r| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Using this representation, we obtain expression for u

$$u(x) = \frac{1}{4\epsilon^2\pi}\int_{R_3}\exp\left(-\frac{|x-y|}{\epsilon}\right)\frac{f(y)}{|x-y|}dy$$

One observes that the smoothness of u is controlled by ϵ. When $\epsilon \to 0$, the kernel $K(r)$ tends to the delta-function, and $u(x) \to f(x)$.

Remark 2.2 Similar explicit solutions can be derived for R_2 and for some bounded domains, such as rectangles or circles (spheres). Considering the approximation problem on a bounded domain, a more efficient way to construct solution is to use the eigenfunction expansion, as it was explained in Section **??**.

Remark 2.3 In contrast with one-dimensional problems, the Green's function is unbounded but integrable.

Example 2.8 (Analytic continuation) A close problem is the analytic continuation. Let $\Omega \subset R_2$ be a domain in a plane with a differentiable boundary $\partial\Omega$. Let $\phi(s)$ be a differentiable function of the point s of $\partial\Omega$. Consider the following problem of *analytic continuation*: Find a function $u(x)$ in Ω such that it coincides with ϕ on the boundary, $u(s) = \phi(s), \forall s \in \partial\Omega$ and minimizes the integral of $(\nabla u)^2$ over Ω. Thus, we formulate a variational problem:

$$\min \int_\Omega (\nabla u)^2 dx \ \text{ in } \Omega, \quad \text{subject to} \quad u|_{\partial\Omega} = \phi$$

Compute the stationarity conditions. We have

$$\frac{\partial(\nabla u)^2}{\partial \nabla u} = 2\nabla u, \quad \nabla\cdot\frac{\partial(\nabla u)^2}{\partial \nabla u} = 2\nabla\cdot\nabla u = 2\Delta u = 0$$

which demonstrates that the minimizer must be harmonic in Ω or be a real part of an analytic function. This explains the name "analytic continuation".

Remark 2.4 Notice that the one-dimensional case is trivial: Ω is an interval, the boundary consists of two points, the minimizer is a straight line between these points. In this sense, harmonic functions are two-dimensional (or higher-dimensional) generalization of linear functions.

2.4 Change of coordinates

In order to transform the variational conditions to polar, spherical, or other coordinates, consider the transformation of the independent variables $x = w(\xi)$ in a multivariable variational problem. Consider the Jacobian of the transformation J, J being the matrix with the elements $J_{ij} = \{\frac{\partial w_i}{\partial \xi_j}\}$, and assume that $\det(J)$ is not zero in all points of Ω. In the new variables, the domain Ω becomes Ω_ξ, the differential dx is transformed as

$$dx = \det(J) d\xi$$

. By the chain rule, gradient $\nabla_x u$ in x coordinates becomes

$$\nabla_x u = \nabla_\xi u \frac{\partial \xi}{\partial x} = J^{-1} \nabla_\xi u$$

where ∇_ξ is the gradient in ξ-coordinates.

The integral

$$R = \int_\Omega F(x, u, \nabla u) dx$$

becomes

$$R = \int_{\Omega_\xi} F_\xi(\xi, u, \nabla_\xi u) d\xi$$

where F_ξ is defined as follows

$$F_\xi(\xi, u, \nabla_\xi u) = F(w(\xi), u, J^{-1}(\xi) \nabla_\xi u) \det J(\xi) \tag{26}$$

The Euler equation in the ξ-coordinates becomes $S_{F_\xi}(u) = 0$, where

$$S_{F_\xi}(u) = \frac{\partial F_\xi}{\partial u} - \nabla_\xi \cdot \frac{\partial F_\xi}{\partial \nabla_\xi u} \tag{27}$$

and the derivatives are related as

$$\frac{\partial F_\xi}{\partial u} = (\det J) \frac{\partial F}{\partial u} \quad \text{and} \quad \frac{\partial F_\xi}{\partial \nabla_\xi u} = (\det J) J^{-1} \frac{\partial F}{\partial \nabla u}$$

Example 2.9 (Helmholtz equation in polar coordinates) Let F be the Lagrangian corresponding to the Helmholtz equation on the plane with Cartesian coordinates (x, y):

$$F = |\nabla u|^2 + \alpha u^2 = u_x^2 + u_y^2 + \alpha u^2$$

We transform it to the polar coordinates (r, θ) using $x = r\cos\theta$, $y = r\sin(\theta)$ and compute the Euler equation for F. We have

$$J = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}, \quad \det J = r$$

Then

$$F_\xi = F_\xi(r, \theta, u, \nabla_\xi u) = r\left[\left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2 + \alpha\, u^2\right]$$

and the Euler equation for the problem (the Helmholtz equation) becomes

$$\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r}\frac{\partial^2 u}{\partial \theta^2} - \alpha r u = 0$$

or, in a more conventional form,

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} - \alpha u = 0$$

2.5 First integrals

Independence of the gradient of minimizer If the Lagrangian is independent of ∇u, $F = F(x, u)$, the Euler equation becomes an algebraic relation

$$\frac{\partial F}{\partial u} = 0.$$

As in one-dimensional case, the minimizer u that solves this equation does not need to be differentiable, even continuous function of x.

Independence of the minimizer If the Lagrangian is independent of u, $F = F(x, \nabla u)$ then Euler equation becomes

$$\nabla \cdot \left(\frac{\partial F}{\partial \nabla u}\right) = 0.$$

Instead of the constancy of $\frac{\partial F}{\partial u'}$ in one-dimensional case, here we state only the divergencefree nature of $\frac{\partial F}{\partial \nabla u}$. Any divergence-free vector admit the representation through curl of a vector potential.

$$\frac{\partial F}{\partial \nabla u} = \nabla \times \psi \tag{28}$$

In the one-variable case, $\nabla \times \psi$ is replaced by a constant and we obtain the first integral; in multivariable case, no additional first integrals exist.

Example 2.10 The Lagrangian $F_1 = \left(\frac{du}{dt}\right)^2, \ t \in R^1$ is a one-dimensional analog of the two-dimensional Lagrangian $F_2 = |\nabla u|^2$. The Euler equation for the one-dimensional problem with this Lagrangian $\frac{d}{dt}\frac{\partial L}{\partial u'} - \frac{\partial L}{\partial u}$, where $u' = \frac{du}{dt}$, has the first integral

$$C_1 = \frac{\partial F_1}{\partial u'} = \frac{du}{dt}$$

that has a solution $u = C_1 t + C_2$.

In multivariable case, for $F_2 = |\nabla u|^2$ we compute $\frac{\partial F}{\partial \nabla u} = 2\nabla u = V$. Here, we denote the gradient by $V = (v_1, v_2)$, $V = \nabla u$. The stationarity condition $\nabla \cdot V = 0$ or

$$\frac{\partial}{\partial x_1} v_1 + \frac{\partial}{\partial x_2} v_2 = 0$$

are identically satisfied if v admits the representation

$$v_1 = \frac{\partial \psi}{\partial x_2} \quad \text{and} \quad v_2 = -\frac{\partial \psi}{\partial x_1}$$

where $\psi(x_1, x_2)$ is an arbitrary potential, that is if (28) holds. Function ψ is called the dual potential, see below, Section **??**. Instead of being a linear function as in one-dimensional case, the the minimizer u is harmonic – a solution to the Laplace equation $\Delta u = 0$. Potential ψ is a conjugate harmonic function.