Euler equations for multiple integrals

January 22, 2013

Contents

1 Reminder of multivariable calculus **2**

1.1 Vector differentiation 2

1.2 Matrix differentiation 3

1.3 Multidimensional integration 7

This part deals with multivariable variational problems that describe equilibria and dynamics of continua, optimization of shapes, etc. We consideration of a multivariable domain $x \in \Omega \subset R_d$ and coordinates $x = (x_1, \dots, x_d)$. The variational problem requires minimization of an integral over Ω of a Lagrangian $L(x, u, Du)$ where $u = u(x)$ is a multivariable vector function and ∇u of minimizers and matrix $Du = \nabla u$ of gradients of these minimizers.

The optimality conditions for these problems include 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 Reminder of multivariable calculus

1.1 Vector differentiation

We remind the definition of 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 $a = (a_1, \dots, a_n)^T$, then the derivative $\frac{d\phi}{da}$ is a row vector

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

assuming that all partial derivatives exist.

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

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

Indeed, the left-hand side is a scalar and the second multiplier in the right-hand side is a column vector, therefore the first multiplier 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 $\phi(a) = |a|^2 = a_1^2 + \dots a_n^2$, the derivative is
$$\frac{d}{da}|a|^2 = 2a^T$$
2. The vector derivative of the euclidean norm $|a|$ of a nonzero 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:

$$\frac{d}{da} c^T a = c$$

Similarly, if C is a $k \times n$ matrix, derivative of a product Ca equals C^T,

$$\frac{d}{da} Ca = C^T$$

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: $\phi_\nu = \nabla\phi \cdot \nu$. 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 vector If $u = (u_1, \ldots, u_n)$ is a vector function, $u_j = u_j(x_1, \ldots x_d)$, $x \in R^d$, 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,

$$\nabla u = Du = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_2}{\partial x_1} & \ldots & \frac{\partial u_n}{\partial x_1} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial u_1}{\partial x_d} & \frac{\partial u_2}{\partial x_d} & \ldots & \frac{\partial u_n}{\partial x_d} \end{pmatrix} = (\nabla u_1 | \nabla u_2 | \ldots | \nabla u_n) \tag{3}$$

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

$$A : B = \operatorname{Tr}(A\,B) \tag{4}$$

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

$$\frac{d\phi}{dA} = \begin{pmatrix} \frac{\partial \phi}{\partial a_{11}} & \frac{\partial \phi}{\partial a_{21}} & \dots & \frac{\partial \phi}{\partial a_{m1}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \phi}{\partial a_{1n}} & \frac{\partial \phi}{\partial a_{2n}} & \dots & \frac{\partial \phi}{\partial a_{mn}} \end{pmatrix} \tag{5}$$

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,

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

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:
$$\frac{d}{dA}\operatorname{Tr} A^{-1} = -A^{-2}.$$
6. Similarly, we compute the derivative of the quadratic form associated with the inverse matrix:
$$\frac{d}{dA} x^T A^{-1} x = -x A^{-2} x^T.$$

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 Newton-like or 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 $\det J$ is the Jacobian and J is the $d\times d$ 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

$$I = \int_{\Omega_\xi} f(x(\xi))(\det \nabla_\xi x)\,dx \tag{6}$$

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

$$\int_\Omega \nabla \cdot a \, dx = \int_{\partial\Omega} a \cdot \nu \, ds \tag{7}$$

Here, ν is the outer normal to Ω and $\nabla\cdot$ is the *divergence* operator,

$$\nabla \cdot a = \frac{\partial a_1}{\partial x_1} + \ldots, + \frac{\partial a_n}{\partial x_n}.$$

If $d = 1$, domain Ω becomes an interval $[c, d]$, the normal show the direction along this interval, and (7) becomes

$$\int_c^d \frac{da}{dx} dx = a(d) - a(c).$$

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

$$\int_\Omega (a \cdot \nabla b) \, dx = -\int_\Omega (b \nabla \cdot a) \, dx + \int_{\partial\Omega} (a \cdot \nu) b \, ds \tag{8}$$

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 two differentiable in Ω vector fields a and b:

$$\int_\Omega (a \cdot \nabla \times b) \, dx = \int_\Omega (b \cdot \nabla \times a) \, dx - \int_{\partial\Omega} (a \times b \cdot \nu) \, ds \tag{9}$$

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

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