Chapter 6

Vector-Valued Minimizers

In this chapter we study minimization problems in which the unknown function y takes values in $\mathbb{R}^n$. As one would expect, solutions of such problems must satisfy appropriate analogues of the first and second Euler-Lagrange equations. It is interesting to note that although the first Euler-Lagrange equation will actually be a system of n (scalar) differential equations for the n unknown components of y, the second Euler-Lagrange will be a single (scalar) equation. Consequently, unless $n = 1$, the second equation will not provide enough information to completely determine solutions to minimization problems.

The treatment of boundary conditions will be a very important issue. In addition to cases where the values of the admissible functions y are either completely prescribed or completely free at each endpoint, we can consider situations where some components of y are prescribed at an endpoint, but the other components are left free. There are other important types of boundary conditions as we shall see below.

6.1 Basic Theory

Let $a, b \in \mathbb{R}$ with $a < b$, $n \in \mathbb{N}$ and $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given. Assume that f is continuously differentiable. We will take $\mathfrak{X} := C^1([a, b]; \mathbb{R}^n)$ as our underlying linear space. The functions in $\mathfrak{X}$ map $[a, b]$ into $\mathbb{R}^n$. We consider the problem of minimizing functionals of the form

$$J(y) = \int_a^b f(x, y(x), y'(x))\, dx$$

over a subset of $\mathfrak{X}$.

As mentioned above, in addition to problems in which the values of y are either completely prescribed or completely free at each endpoint, there are many other important possibilities. For example, with $n = 3$, we may wish to consider boundary conditions such as

$$y_1(a) + 2y_2(a) - y_3(a) = 5 \text{ and } y_1(a) + 3y_3(a) = -4. \tag{6.1}$$

Notice that (6.1) can be viewed as a pair of linear equations in the three variables $y_1(a), y_2(a), y_3(a)$. In general, we should not prescribe more than n such equations at an endpoint because otherwise the system of linear equations would either be inconsistent or redundant. If fewer than n equations are prescribed, we can always make them into exactly n equations by augmenting the system with $0 = 0$ as many times as necessary.

Therefore, we consider boundary conditions of the form

$$\mathcal{A}y(a)^\mathsf{T} = \xi^\mathsf{T} \text{ and } \mathcal{B}y(b)(b)^\mathsf{T} = \eta^\mathsf{T}, \tag{6.2}$$

with $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n\times n}$ and $\xi, \eta \in \mathbb{R}^n$. Notice that if $y(a)$ is completely prescribed then an appropriate choice for $\mathcal{A}$ would be $\mathcal{A} = I$, where I is the $n \times n$ identity matrix and we would simply take ξ to be the value prescribed for $y(a)$. (Observe that any invertible matrix $\mathcal{A}$ would also work – but with a different choice of ξ.) For a problem with a completely free end at b, an appropriate choice for $\mathcal{B}$ and η would be $\mathcal{B} = 0, \eta = 0$. Then

$$\mathcal{B}y(b)^\mathsf{T} = 0^\mathsf{T} = \eta^\mathsf{T}$$

regardless of what $y(b)$ is.

As a final illustration, suppose that $n = 3$ and we want the boundary conditions expressed in (6.1) Then we can take

$$\mathcal{A} = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } \xi^\mathsf{T} = \begin{pmatrix} 5 \\ -4 \\ 0 \end{pmatrix}.$$

In anticipation of computing the Gateaux variations of J, we extend the notation discussed in Section 2.1. We still denote the partial derivative of f with respect to its first argument by $f_{,1}(x, y, z)$. However, since the second and third arguments of f are vectors, it is appropriate to introduce *partial gradients*. We use $f_{,2}(x, y, z)$ to denote the partial gradient of f with respect to the components of the second argument. Similarly, $f_{,3}(x, y, z)$ denotes the partial gradient of f with respect to the components of the third argument. As an example, suppose that $n = 3$ and

$$f(x, y, z) = x^2y_1^2 + x^3y_2y_3 + z_1^2 + z_3^2 + y_1z_2 \quad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R}^3 \times \mathbb{R}^3.$$

Then

$$\begin{aligned} f_{,1}(x, y, z) &= 2xy_1^2 + 3x^2y_2y_3 && \text{for all } (x, y, z) \in [a, b] \times \mathbb{R}^3 \times \mathbb{R}^3, \\ f_{,2}(x, y, z) &= (2xy_1 + z_2, x^3y_3, x^3y_2) && \text{for all } (x, y, z) \in [a, b] \times \mathbb{R}^3 \times \mathbb{R}^3 \end{aligned}$$

and

$$f_{,3}(x, y, z) = (2z_1, y_1, 2z_3) \qquad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R}^3 \times \mathbb{R}^3.$$

We now summarize our problem. Let $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n\times n}$ and $\xi, \eta \in \mathbb{R}^n$ be given and put

$$\mathscr{Y} := \left\{ y \in C^1([a,b];\mathbb{R}^n) \,:\, \mathcal{A}y(a)^\mathsf{T} = \xi^\mathsf{T} \text{ and } \mathcal{B}y(b)^\mathsf{T} = \eta^\mathsf{T} \right\}.$$

We assume that ξ^T is in the range of $\mathcal{A}$ and that η^T is in the range of $\mathcal{B}$ (otherwise $\mathscr{Y}$ is empty). Define $J : \mathscr{Y} \to \mathbb{R}$ by

$$J(y) := \int_a^b f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathscr{Y}.$$

We wish is to minimize J over $\mathscr{Y}$.

We will now derive the analogue of $(\text{E-L})_1$ for J. The space of admissible variations for each $y \in \mathscr{Y}$ is easily seen to be

$$\mathscr{V} := \left\{ v \in C^1([a,b];\mathbb{R}^n) \,:\, \mathcal{A}v(a)^\mathsf{T} = \mathcal{B}v(b)^\mathsf{T} = 0 \right\}.$$

Let $y \in \mathscr{Y}$ and $v \in \mathscr{V}$ be given. A straightforward computation utilizing the chain rule yields

$$\delta J(y;v) = \int_a^b \Big\{ f_{,2}(x, y(x), y'(x)) \cdot v(x) + f_{,3}(x, y(x), y'(x)) \cdot v'(x) \Big\}\, dx.$$

Suppose that $y_* \in \mathscr{Y}$ is a minimizer for J over $\mathscr{Y}$. Then y_* must satisfy $\delta J(y_*; v) = 0$ for every $v \in \mathscr{V}$. To derive the Euler-Lagrange equations for J and the natural boundary conditions we need to use some clever choices for $v \in \mathscr{V}$. Put

$$\mathscr{W} := \left\{ w \in C^1[a,b] \,:\, w(a) = w(b) = 0 \right\}.$$

Observe that the elements of $\mathscr{W}$ are scalar-valued functions. Let $w \in \mathscr{W}$ and $\mu \in \mathbb{R}^n$ be given. Notice that $v := w\mu \in \mathscr{V}$. With this choice for v, we have

$$\delta J(y_*;v) = \int_a^b \Big\{ f_{,2}(x, y_*(x), y_*'(x)) \cdot \mu w(x) + f_{,3}(x, y_*(x), y_*'(x)) \cdot \mu w'(x) \Big\}\, dx.$$

If y_* minimizes J over $\mathscr{Y}$, then the above expression is zero for each $w \in \mathscr{W}$ and $\mu \in \mathbb{R}^n$. Putting $F(x) = f_{,2}(x, y_*(x), y_*'(x)) \cdot \mu$ and $G(x) = f_{,3}(x, y_*(x), y_*'(x)) \cdot \mu$ for $x \in [a,b]$, Lemma 3.4 implies that $G \in C^1[a,b]$ and

$$\frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x)) \cdot \mu \right] = f_{,2}(x, y_*(x), y_*'(x)) \cdot \mu \quad \text{for all } x \in [a,b] \text{ and } \mu \in \mathbb{R}^n.$$

Since, G is continuously differentiable for every choice of $\mu \in \mathbb{R}^n$, we can conclude that the mapping $x \mapsto f_{,3}(x, y_*(x), y_*'(x))$ is continuously differentiable and consequently

$$\left\{ f_{,2}(x, y_*(x), y_*'(x)) - \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x)) \right] \right\} \cdot \mu = 0 \quad \text{for all } x \in [a,b] \text{ and } \mu \in \mathbb{R}^n.$$

By Lemma 2.1, we have

$$f_{,2}(x, y_*(x), y_*'(x)) = \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x))\right] \quad \text{for all } x \in [a,b]. \qquad \text{(E-L)}_1$$

Now, we derive the associated natural boundary conditions. To analyze the endpoint $x = a$, we put $w(x) = \frac{x-b}{a-b}$ for $x \in [a,b]$, so that $w(a) = 1$ and $w(b) = 0$. Suppose that $\lambda \in \mathbb{R}^n$ satisfies

$$\mathcal{A}\lambda^{\mathsf{T}} = 0,$$

i.e. the vector λ^{T} is in the null-space of $\mathcal{A}$, and put $v = w\lambda$. Since $w(b) = 0$ and λ^{T} is in the null-space of $\mathcal{A}$, we find $v \in \mathscr{V}$. With this choice for v, we have

$$\delta J(y_*; v) = \int_a^b \Big\{ f_{,2}(x, y_*(x), y_*'(x)) \cdot \lambda w(x) + f_{,3}(x, y_*(x), y_*'(x)) \cdot \lambda w'(x) \Big\}\, dx.$$

Upon integrating the second term by parts and using the fact that y_* satisfies (E-L)_1 for J, we have

$$\delta J(y_*; v) = f_{,3}(x, y_*(x), y_*'(x)) \cdot \lambda w(x)\Big|_a^b = 0 \Rightarrow f_{,3}(a, y_*(a), y_*'(a)) \cdot \lambda = 0.$$

Thus y_* must satisfy the natural boundary conditions

$$f_{,3}(a, y_*(a), y_*'(a)) \cdot \lambda = 0 \quad \text{for all } \lambda \in \mathbb{R}^n \text{ satisfying } \mathcal{A}\lambda^{\mathsf{T}} = 0. \qquad \text{(NBC)}_a$$

Similarly, the natural boundary conditions at $x = b$ are

$$f_{,3}(b, y_*(b), y_*'(b)) \cdot \lambda = 0 \quad \text{for all } \lambda \in \mathbb{R}^n \text{ satisfying } \mathcal{B}\lambda^{\mathsf{T}} = 0. \qquad \text{(NBC)}_b$$

We have just proved the following

Theorem 6.1 *Let $a, b \in \mathbb{R}$ with $a < b$ be given, and let $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given. Assume that f is continuously differentiable. Let $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n\times n}$ and $\xi, \eta \in \mathbb{R}^n$ be such that ξ^{T} and η^{T} are in the range of $\mathcal{A}$ and $\mathcal{B}$, respectively. Put*

$$\mathscr{Y} := \left\{ y \in C^1([a,b]; \mathbb{R}^n) \ : \ \mathcal{A}y(a)^{\mathsf{T}} = \xi^{\mathsf{T}} \text{ and } \mathcal{B}y(b)^{\mathsf{T}} = \eta^{\mathsf{T}} \right\},$$

and define the functional $J : \mathscr{Y} \to \mathbb{R}$ by

$$J(y) = \int_a^b f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathscr{Y}.$$

Let $y_* \in \mathscr{Y}$ be given and assume that y_* minimizes (or maximizes) J over $\mathscr{Y}$. Then the mapping $x \mapsto f_{,3}(x, y_*(x), y_*'(x))$ is continuously differentiable on $[a,b]$ and y_* must satisfy the system of n differential equations

$$f_{,2}(x, y_*(x), y_*'(x)) = \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x))\right] \quad \text{for all } x \in [a,b], \qquad \text{(E-L)}_1$$

and the natural boundary conditions

$$f_{,3}(a, y_*(a), y_*'(a)) \cdot \lambda = 0 \quad \text{for all } \lambda \in \mathbb{R}^n \textit{ satisfying } \mathcal{A}\lambda^{\mathsf{T}} = 0 \qquad (\mathrm{NBC})_a$$

and

$$f_{,3}(b, y_*(b), y_*'(b)) \cdot \lambda = 0 \quad \text{for all } \lambda \in \mathbb{R}^n \textit{ satisfying } \mathcal{B}\lambda^{\mathsf{T}} = 0. \qquad (\mathrm{NBC})_b$$

Let us make a few remarks.

Remark 6.1 *If for each $x \in [a, b]$ the function $f(x, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is convex, then J is convex and any $y_* \in \mathscr{Y}$ that satisfies (E-L)$_1$, (NBC)$_a$ and (NBC)$_b$ must be a minimizer for J over $\mathscr{Y}$.*

Remark 6.2 *Constraints of the form*

$$\int_a^b g(x, y(x), y'(x))\, dx = c,$$

where $g : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $c \in \mathbb{R}$ are given, can be handled by the method of Lagrange multipliers.

Remark 6.3 *Minimizers for J over $\mathscr{Y}$ must also satisfy an analogue of the second Euler-Lagrange equation. If J attains a minimum over $\mathscr{Y}$ at $y \in \mathscr{Y}$, then y must satisfy the integro-differential equation*

$$f(x, y(x), y'(x)) - y'(x) \cdot f_{,3}(x, y(x), y'(x)) = c + \int_a^x f_{,1}(t, y(t), y'(t))\, dt \quad (\mathrm{E\text{-}L})_2$$

$$\text{for all } x \in [a, b],$$

for some $c \in \mathbb{R}$. Notice that (E-L)$_2$ is a single equation with n unknown functions. The fact that a minimizer satisfies the second Euler-Lagrange equation, though useful, does not provide as much information about the minimizer as the fact that it satisfies the first Euler-Lagrange equation.

6.2 Example 6.2

Let us look at an example with $n = 2$. Put $\mathfrak{X} = C^1([0, 1]; \mathbb{R}^2)$,

$$\mathscr{Y} := \left\{ y \in \mathscr{C}^1([0, 1]; \mathbb{R}^2) \,:\, y_1(0) = y_2(0) = 0 \text{ and } y_1(1) + y_2(1) = 1 \right\},$$

and define the functional $J : \mathscr{Y} \to \mathbb{R}$ by

$$J(y) := \int_0^1 \left[y_1'(x)^2 + y_2'(x)^2 + 2y_1(x)y_2(x) \right] dx \quad \text{for all } y \in \mathscr{Y}.$$

We wish to minimize J over $\mathscr{Y}$.

First, we write the boundary conditions in the form (6.2). Since both components of each admissible y are prescribed to be 0 at $x = 0$, we take $\mathcal{A} := I$ and $\xi := 0$. To write the boundary conditions at $x = 1$ in the form of (6.2), we may take

$$\mathcal{B} := \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } \eta^{\mathsf{T}} := \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

With these definitions, we see that

$$\mathscr{Y} = \left\{ y \in C^1([0,1];\mathbb{R}^2) \,:\, \mathcal{A}y(0)^{\mathsf{T}} = \xi^{\mathsf{T}} \text{ and } \mathcal{B}y(1)^{\mathsf{T}} = \eta^{\mathsf{T}} \right\}.$$

The integrand $f : [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ for J is given

$$f(x,y,z) = z_1^2 + z_2^2 + 2y_1y_2 \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2.$$

Consequently, we have

$$f_{,2}(x,y,z) = (2y_2, 2y_1) \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2, \tag{6.3}$$

and

$$f_{,3}(x,y,z) = (2z_1, 2z_2) \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2. \tag{6.4}$$

At $x = 0$, there are no natural boundary conditions, because the null-space for $\mathcal{A} = I$ is the singleton $\{0\}$. This is what should be expected, since the value of $y \in \mathscr{Y}$ is completely prescribed at $x = 0$. To find the natural boundary conditions at $x = 1$, we need to determine the null-space of $\mathcal{B}$, i.e.we want to find those $\lambda \in \mathbb{R}^2$ such that

$$\mathcal{B}\lambda^{\mathsf{T}} = 0.$$

Using our definition for $\mathcal{B}$, we require

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0.$$

Thus, the null-space for $\mathcal{B}$ consists of those $\lambda \in \mathbb{R}^2$ satisfying

$$\lambda_2 = -\lambda_1.$$

Using this and (6.4), the natural boundary conditions at $x = 1$ are

$$(2y_1'(1), 2y_2'(1)) \cdot (\lambda_1, -\lambda_1) = 0, \quad \text{for all } \lambda_1 \in \mathbb{R}. \tag{NBC$_1$}$$

or equivalently

$$y_1'(1) = y_2'(1). \tag{NBC$_1$}$$

We now write down the first Euler-Lagrange equations. From (6.3) and (6.4), we have

$$\begin{aligned} & f_{,2}(x, y(x), y'(x)) = \frac{d}{dx}\left[f_{,3}(x, y(x), y'(x))\right] && (\text{E-L})_1 \\ \Rightarrow & (2y_2(x), 2y_1(x)) = \frac{d}{dx}\left[(2y_1'(x), 2y_2'(x))\right] \\ \Rightarrow & (2y_2(x), 2y_1(x)) = (2y_1''(x), 2y_2''(x)) \\ \Rightarrow & \left\{ \begin{array}{l} y_1''(x) = y_2(x) \\ y_2''(x) = y_1(x) \end{array} \right. \quad \text{for all } x \in [0,1]. \end{aligned}$$

To find the general solution to $(\text{E-L})_1$, observe that since $y \in \mathscr{Y} \subset C^1([0,1];\mathbb{R}^2)$, if y satisfies the above equations, then y must have continuous derivatives of all orders. Thus

$$\left\{ \begin{array}{l} y_1''(x) = y_2(x) \\ y_2''(x) = y_1(x) \end{array} \right. \Rightarrow y_1^{(4)}(x) = y_2''(x) \Rightarrow y_1^{(4)}(x) - y_1(x) = 0 \quad \text{for all } x \in [0,1].$$

The roots for the characteristic equation $r^4 - 1 = 0$ are $r = \pm i, \pm 1$. It follows that y_1 must have the form

$$y_1(x) = c_1 \cosh x + c_2 \sinh x + c_3 \cos x + c_4 \sin x. \tag{6.5}$$

For the second component, we have

$$\begin{aligned} & y_2(x) = y_1''(x) \\ \Rightarrow & y_2(x) = c_1 \cosh x + c_2 \sinh x - c_3 \cos x - c_4 \sin x. \end{aligned} \tag{6.6}$$

Imposing the boundary condition $y_1(0) = y_2(0) = 0$ leads to

$$\left\{ \begin{array}{l} y_1(0) = c_1 + c_3 = 0 \\ y_2(0) = c_1 - c_3 = 0 \end{array} \right. \Rightarrow c_1 = 0 \text{ and } c_3 = 0.$$

Substituting these values for c_1 and c_3 into (6.5) and (6.6) yields

$$y_1(x) = c_2 \sinh x + c_4 \sin x \quad \text{and} \quad y_2(x) = c_2 \sinh x - c_4 \sin x. \tag{6.7}$$

Let us now impose the boundary condition $y_1(1) + y_2(1) = 1$. From (6.7), we need

$$c_2 \sinh 1 + c_4 \sin 1 + c_2 \sinh 1 - c_4 \sin 1 = 1.$$

Thus

$$c_2 = \frac{1}{2 \sinh 1},$$

and we now have that

$$y_1(x) = \frac{\sinh x}{2 \sinh 1} + c_4 \sin x \quad \text{and} \quad y_2(x) = \frac{\sinh x}{2 \sinh 1} - c_4 \sin x. \tag{6.8}$$

Lastly, we impose the natural boundary condition $y_1'(1) - y_2'(1) = 0$. Using (6.8), we have

$$y_1'(1) - y_2'(1) = \frac{\cosh 1}{2\sinh 1} + c_4 \cos 1 - \frac{\cosh 1}{2\sinh 1} + c_4 \cos 1 = 2c_4 \cos 1 = 0 \Rightarrow c_4 = 0.$$

Thus, the only possible minimizer for J over $\mathscr{Y}$ is

$$y(x) = (y_1(x), y_2(x)) = \left(\frac{\sinh x}{2\sinh 1}, \frac{\sinh x}{2\sinh 1} \right) \quad \text{for all } x \in [0, 1].$$

6.3 Lagrangian Constraints

Let $a, b \in \mathbb{R}$ with $a < b$ and $n \in \mathbb{N}$ with $n \geq 2$ be given. In this section we consider problems in which the admissible functions $y : [a, b] \to \mathbb{R}^n$ are required to satisfy the constraint

$$g(y(x)) = 0 \quad \text{for all } x \in [a, b],$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a given smooth function satisfying

$$\nabla g(z) \neq 0 \quad \text{for all } z \in \mathbb{R}^n \text{ with } g(z) = 0. \tag{6.9}$$

In other words, we are limiting our attention to admissible functions whose ranges are subsets of the 0-level surface of g.

For simplicity, we shall only treat the case where y is completely prescribed at both endpoints. Let $\xi, \eta \in \mathbb{R}^n$ be given and put

$$\mathscr{Y} = \{ y \in C^1([a, b]; \mathbb{R}^n) \,:\, y(a) = \xi, y(b) = \eta \text{ and } g(y(x)) = 0 \quad \text{for all } x \in [a, b] \}.$$

Let $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given. For technical reasons, we assume that f is twice continuously differentiable and that g is twice continuously differentiable. We shall also assume twice continuous differentiability of our candidate for a minimizer.

We consider the functional $J : \mathscr{Y} \to \mathbb{R}$ defined by

$$J(y) = \int_a^b f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathscr{Y}.$$

Theorem 6.2 *Let $y_* \in \mathscr{Y} \cap C^2([a, b]; \mathbb{R}^n)$ be given and assume that y_* minimizes J on $\mathscr{Y}$. Then there exists a function $\lambda \in C[a, b]$ such that*

$$f_{,2}(x, y_*(x), y_*(x)) - \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x)) \right] = \lambda(x) \nabla g(y_*(x)) \quad \text{for all } x \in [a, b].$$

We shall prove the theorem only in the special case when the constraining function g has the form

$$g(z) = z_n - \psi(z_1, z_2, \cdots, z_{n-1}) \quad \text{for all } z \in \mathbb{R}^n, \tag{6.10}$$

for some function $\psi \in C^2(\mathbb{R}^{n-1})$. The idea is that the first $n-1$ components of y will yield a minimizer for a standard variational problem. The first Euler-Lagrange equation for the "reduced" problem will yield the desired conclusion. (DETAILS TO BE FILLED IN)