Chapter 5

Optimization of Functionals

There are optimization problem that cannot be formulated as minimum norm problems. This chapter discusses the theory and geometry of the more general problems.

1. The concepts of differentials, gradients, and so on can be generalized to normed spaces.
2. Variational theory of optimization is much in parallel to the familiar theory in finite dimensions.

Topics to be discussed include:

- Differentials
- Extrema
- Euler-Lagrange Equations

Differentials

There are several ways to introduce the notion of derivative in normed spaces.

- Gateaux derivative
- Frèchet derivatie

Gauteaux Differential

- Let $T : D \longrightarrow Y$ be a mapping where $D \subset X$, X is a vector space, and Y is a normed space. Given $\mathbf{x} \in D$ and $\mathbf{h} \in X$,
  - ◇ The limit
  $$\delta T(\mathbf{x}; \mathbf{h}) = \lim_{\alpha \to 0} \frac{T(\mathbf{x} + \alpha \mathbf{h}) - T(\mathbf{x})}{\alpha},$$
  if exists, is called the Gateaux differential of T at $\mathbf{x}$ with increment $\mathbf{h}$.
    - ▷ If $\delta T(\mathbf{x}; \mathbf{h})$ exists for each $\mathbf{h} \in X$, the mapping T is said to be Gateaux differentiable at $\mathbf{x}$.
- The limit is takin in the normed space Y.
- The Gauteaux differential is analogous to the directional derivative in that if $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuously differentiable, then
$$\delta f(\mathbf{x}; \mathbf{h}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i = \langle \mathbf{h}, \nabla f(\mathbf{x}) \rangle.$$
- For each fixed $\mathbf{x} \in D$, $\delta T(\mathbf{x}; \cdot) : X \longrightarrow Y$ is a linear map in $\mathbf{h}$.

Fréchet Differential

- Let $T : D \longrightarrow Y$ be a mapping where $D \subset X$ is an open domain and X, Y are normed spaces. Given $\mathbf{x} \in D$, if for each $\mathbf{h} \in X$, the Gauteaux differential $\delta T(\mathbf{x}; \mathbf{h})$ exists, is continuous, and satisfies

$$\lim_{\|\mathbf{h}\| \to 0} \frac{\|T(x+h) - T(x) - \delta T(\mathbf{x}; \mathbf{h})\|}{\|\mathbf{h}\|} = 0,$$

  then T is said to be Frèchet differentiable at $\mathbf{x}$.

  - ◇ At a fixed point $\mathbf{x} \in D$, the bounded linear operator $A_{\mathbf{x}} \in B(X, Y)$ such that

$$A_{\mathbf{x}}\mathbf{h} = \delta T(\mathbf{x}; \mathbf{h})$$

    is called the Fréchet derivative of T at $\mathbf{x}$ and is denoted as $T'(\mathbf{x})$.

- The mapping from X to $B(X, Y)$ via $\mathbf{x} \mapsto T'(\mathbf{x})$ is called the *Fréchet derivative* of T.
- If T is a functional, then $T'(\mathbf{x}) \in X^*$ is called the *gradient* of T at $\mathbf{x}$.
- Much of the theory of ordinary derivatives, including the implicit function theorem and the Taylor series expansion, can be generalized to Fréchet derivatives. (How?)

- Some examples:
  - ◇ If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, then

$$f'(\mathbf{x}) = \left[\frac{\partial f_i(\mathbf{x})}{\partial x_j}\right]$$

where the Jacobian matrix acts like the matrix representation of the linear operator. (Recall how the dual space is represented!)

  - ◇ Suppose $I : C[0,1] \longrightarrow \mathbb{R}$ is defined by

$$I(f) = \int_0^1 g(f(t), t)dt,$$

where $g : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ and $\frac{\partial g}{\partial x}$ exists. Then the action of $I'(f)$ on $h(t) \in C[0,1]$ is given by

$$I'(f)h(t) = \int_0^1 \frac{\partial g(f(t), t)}{\partial x} h(t)dt.$$

    - ▷ May we say that

$$\nabla I(f) = \frac{\partial g(f(t), t)}{\partial x}?$$

      - · Recall that the dual space of $C[0,1]$ is identified with $NBV[0,1]$.

Conditions at Extrema

- Using the norm to define a neighborhood N of a given point $\mathbf{x}_0$ in a normed space, $\mathbf{x}_0$ is said to be a *relative minimum* of a functional f on a subset Ω if $f(\mathbf{x}_0) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \Omega \bigcap N$.

- Suppose the real-valued functional f is Gauteaux differentiable at $\mathbf{x}_0$ in a normed space. A necessary condition for f to have extremum at $\mathbf{x}_0$ is that

$$\delta f(\mathbf{x}_0; \mathbf{h}) = 0$$

  for all $\mathbf{h} \in X$.

  ◇ Simple consequence by the ordinary calculus, but significant impact.

Euler-Lagurange Equations

The Euler-Lagurange equation is

- A direct result of differential calculus applied to the settings of functions.
- The equation usually ends up with a differential equation that needs to be solved further.

Fixed-end Problem

- Find a function $x(t)$ over the interval $[t_1, t_2]$ that minimizes the integral functional

$$J(x) = \int_{t_1}^{t_2} f(x(t), \dot{x}(t), t)dt$$

where f is continuous and has continuous partial derivatives with respect to x and $\dot{x}$.

  - ◇ We must agree on the class $D[t_1, t_2]$ of functions within which we seek the optimal values — the so called *admissible set*. (What class of functions should be considered?)
  - ◇ One version is to assume that $x(t_1)$ and $x(t_2)$ are fixed.
    - ▷ Given $x(t)$ admissible, if $x(t) + h(t)$ is also admissible, then $h(t_1) = h(t_2) = 0$.

- The necessary condition for the extremum is that for all h,

$$\begin{aligned}\delta J(x; h) &= \int_{t_1}^{t_2} \left(f_x(x, \dot{x}, t)h(t) + f_{\dot{x}}(x, \dot{x}, t)\dot{h}(t) \right) dt \\ &= \int_{t_1}^{t_2} \left(f_x(x, \dot{x}, t) - \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t) \right) h(t)dt + f_{\dot{x}}(x, \dot{x}, t)h(t)|_{t_1}^{t_2} = 0.\end{aligned}$$

  - ◇ To use the integration by parts, ... we have to assume that $\frac{d}{dt} f_{\dot{x}}$ is continuous.

- If $\alpha(t)$ and $\beta(t)$ are continuous in $[t_1, t_2]$ and

$$\int_{t_1}^{t_2} \Big(\alpha(t)h(t) + \beta(t)\dot{h}(t)\Big)\, dt = 0$$

  for every $h \in D[t_1, t_2]$ with $h(t_1) = h(t_2) = 0$, then β is differentiable and $\dot{\beta}(t) = \alpha(t)$ in $[t_1, t_2]$.

  ◇ Define $\gamma(t) := \int_{t_1}^{t} \alpha(\tau) d\tau$, then

$$\int_{t_1}^{t_2} \Big(\alpha(t)h(t) + \beta(t)\dot{h}(t)\Big)\, dt = -\int_{t_1}^{t_2} \underbrace{(\gamma(t) + \beta(t))}_{\eta(t)} \dot{h}(t) dt.$$

  ◇ $\eta(t) \equiv c$.

  ▷ Let c be the average value of $\eta(t)$ over $[t_1, t_2]$.

  ▷ Define $k(t) := \int_{t_1}^{t} (\eta(\tau) - c)\, d\tau$ (as a special $h \in D[t_1, t_2]$.

  ▷ Observe that

$$\int_{t_1}^{t_2} (\eta(t) - c)^2\, d\tau = \int_{t_1}^{t_2} (\eta(t) - c)\, \dot{k}(t) dt = \int_{t_1}^{t_2} \eta(t)\dot{k}(t) dt - c[h(t_2) - h(t_1)] = 0.$$

  ◇ $\beta(t) = \gamma(t) + c$ is differentiable.

- At the extremum point, $x(t)$ must satisfy the Euler-Lagrange equation

$$f_x(x, \dot{x}, t) = \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t).$$

Examples

- Assumptions:
  - That a retired professor has no income other than that obtained from his fixed quantity of savings S. (Oh, that is how the government treats us!)
  - That his rate of enjoyment at a given time is $U(r(t))$ where $r(t)$ is his rate of expenditure. (Uhmmm, life is more than just material things!)
  - That the true sense of enjoyment is degraded in time as $e^{-\beta t}U(r(t))$. (Alas, money cannot buy all the joy!)
  - That the current capital $x(t)$ at time t generates an investment interest $\alpha x(t)$. (So this is the only bright thing in this scenario!)
  - That the man knows exactly that he will live for a period of time $[0, T]$ when he begins to withdraw his savings. (Attention! Do not think that you can make this prediction!)
- Objective:

$$\begin{aligned} \text{Maximize} \quad & \int_0^T e^{-\beta t} U(r(t))dt \\ \text{Subject to} \quad & x(0) = S, \quad x(T) = 0, \end{aligned}$$

whereas

$$\dot{x}(t) = \alpha x(t) - r(t).$$

  - What is the lifetime plan of investment and expenditure that maximize the total enjoyment?

- Optimal solution $x(t)$ must satisfy:

$$\alpha e^{-\beta t} U'(\alpha x(t) - \dot{x}(t)) + \frac{d}{dt} e^{-\beta t} U'(\alpha x(t) - \dot{x}(t)) = 0.$$

  ◇ Equivalently,

$$\begin{aligned} \frac{d}{dt} U'(r(t)) &= (\beta - \alpha) U'(r(t)), \\ U'(r(t)) &= U'(r(0)) e^{(\beta - \alpha)t}. \end{aligned}$$

  ◇ Once U is specified, $x(t)$ and hence $r(t)$ can be completely characterized from the boundary conditions $x(0) = x(T) = 0$.

Equality Constrained Problems

- Fixed-end problems are a kind of constrained problem. Since these kinds of constrained are explicitly described, they can be embedded in the space of admissible solutions.
- For implicitly defined constraints, the theory of Lagrange multiplier is quite useful.

$$\begin{array}{ll}\text{Minimize} & f(\mathbf{x}),\\ \text{Subject to} & g_i(\mathbf{x}) = 0, \quad i = 1, \ldots, n.\end{array}$$

  - ◇ Recall that there is a geometric meaning of the Lagrange multiplier in finite dimensional space. Such a concept needs to be generalized in normed vector space.
  - ◇ The Lagrange multiplier theory usually ends up with a new functional whose stationary points can be found by the Euler-Lagrange equation.

Notion of Tangent in Normed Space

- Assume
  - ◇ That $\mathbf{x}_0$ is an extremum of the functional f subject to the functional constraints $g(\mathbf{x}) = 0$, $i = 1, \ldots n$.
  - ◇ That the linear functionals $g_1'(\mathbf{x}_0), \ldots, g_n'(\mathbf{x}_0)$ are linear independent.
- Then

$$\delta g_i(\mathbf{x}_0; \mathbf{h}) = 0, \quad i = 1, \ldots, n \Rightarrow \delta f(\mathbf{x}_0; \mathbf{h}) = 0.$$

  - ◇ Note that the concept of tangent is built in the notion of Gateaux differential of g_i at $\mathbf{x}_0$ with arbitrary increment $\mathbf{h}$.

- Here is the proof!
  - There exist n linearly independent vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n \in X$, such that the matrix

$$M := [\delta g_i(\mathbf{x}_0; \mathbf{y}_j)]$$

    is nonsingular.
  - Consider the system of equations from $\mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$ defined by

$$\begin{aligned} g_1(\mathbf{x}_0 + \alpha \mathbf{h} + \sum_{i=1}^n \beta_i \mathbf{y}_i) &= 0, \\ g_2(\mathbf{x}_0 + \alpha \mathbf{h} + \sum_{i=1}^n \beta_i \mathbf{y}_i) &= 0, \\ &\vdots \\ g_n(\mathbf{x}_0 + \alpha \mathbf{h} + \sum_{i=1}^n \beta_i \mathbf{y}_i) &= 0, \end{aligned}$$

    for variables $(\alpha, \beta_1, \ldots \beta_n)$.
  - With a little bit abuse of the notation, the Jacobian

$$\left[\frac{\partial g_i}{\partial \beta_j}\right|_{\alpha=0, \boldsymbol{\beta}=\mathbf{0}} = M.$$

    is nonsingular.

- ◇ The implicit function theorem (in $\mathbb{R}^n$) kicks in.
  - ▷ There exists a function $\boldsymbol{\beta}(\alpha)$ in the neighborhood of $\alpha = 0$ such

$$\begin{aligned} 0 &= g_i(\mathbf{x}_0 + \alpha\mathbf{h} + \sum_{j=1}^{n} \beta_j(\alpha)\mathbf{y}_j) \\ &= \underbrace{g_i(\mathbf{x}_0) + \alpha\delta g_i(\mathbf{x}_0;\mathbf{h})}_{0} + \underbrace{\delta g_i(\mathbf{x}_0;\sum_{j=1}^{n}\beta_j(\alpha)\mathbf{y}_j)}_{M\boldsymbol{\beta}(\alpha)} + o(\alpha) + o(\|\sum_{j=1}^{n}\beta_j(\alpha)\mathbf{y}_j)\|). \end{aligned}$$

  - ▷ But

$$M\boldsymbol{\beta}(\alpha) = O(\|\sum_{j=1}^{n}\beta_j(\alpha)\mathbf{y}_j)\|) \Rightarrow \|\sum_{j=1}^{n}\beta_j(\alpha)\mathbf{y}_j)\| = o(\alpha).$$

- ◇ Along the one-parameter curve $\mathbf{x}_0 + \alpha\mathbf{h} + \sum_{j=1}^{n}\beta_j(\alpha)\mathbf{y}_j$, the functional f assume the extremum at $\mathbf{x}_0$. So

$$\frac{d}{d\alpha} f(\mathbf{x}_0 + \alpha\mathbf{h} + \sum_{j=1}^{n}\beta_j(\alpha)\mathbf{y}_j)|_{\alpha=0} = 0.$$

- ◇ The above derivative is precisely $\delta f(\mathbf{x}_0;\mathbf{h}) = 0$.

Lagrange Multiplier

- Let $g_1, \ldots, g_n$ be linear independent linear functionals on a vector space X. Let f be another linear functional on X such that for every $x \in X$ satisfying $g_i(x) = 0$, $i = 1, \ldots, n$, we always have $f(x) = 0$. Show that there are constants $\lambda_1, \ldots, \lambda_n$ such that

$$f = \sum_{i=1}^{n} \lambda_i g_i.$$

- Assume that
  - ◇ The point $\mathbf{x}_0$ is an extremum of the functional f subject to the constraints $g_i(\mathbf{x}) = 0$, $i = 1, \ldots, n$.
  - ◇ The linear functionals $g_1'(\mathbf{x}_0), \ldots, g_n'(\mathbf{x}_0)$ are linear independent.

  Then there exists constants $\lambda_1, \ldots, \lambda_n$ such that

$$\delta f(\mathbf{x}_0; \mathbf{h}) + \sum_{j=1}^{n} \lambda_j \delta g_j(\mathbf{x}_0; \mathbf{h}) = 0$$

  for every $\mathbf{h} \in X$.

Isometric Problem

- Find the curve with length ℓ and end points $(-1, 0)$ and $(1, 0)$ so that it encloses maximum area between the curve and the x-axis.
- Mathematical model:

$$\begin{aligned} \text{Maximize} \quad & \int_{-1}^{1} y(x)dx, \\ \text{Subject to} \quad & \int_{-1}^{1} \sqrt{1 + (y'(x))^2}dx = \ell, \\ & y(-1) = y(1) = 0. \end{aligned}$$

- By the Lagrange multiplier,

$$\delta J(y; h) = 0$$

  for every admissible h where the functional J is defined by

$$J(y) = \int_{-1}^{1} (y + \lambda\sqrt{1 + (y'(x))^2})dx.$$

- Apply the Euler-Lagrange equation,

$$1 - \lambda \frac{d}{dx} \frac{y'(x)}{\sqrt{1 + (y'(x))^2}} = 0.$$

  ◇ Show that the solution is the arc of a circle.