Introduction to Calculus of Variations

In calculus we search for the solution to

$$\min_{x \in \mathcal{X}} \; f(x) \qquad\qquad [*]$$

where $\mathcal{X}$ is a finite-dimensional space, *e.g.*, $\mathcal{X} \subseteq \mathbb{R}^n$.

If $n = 1$ and $\mathcal{X} = [a, b]$, then under some smoothness conditions we can characterize solutions to $[*]$ through a set of *necessary conditions*.

Necessary conditions for a minimum at x^*:

Interior point: $f'(x^*) = 0$, $f''(x^*) \geq 0$, and $a < x^* < b$.
Left Boundary: $f'(x^*) \geq 0$ and $x^* = a$.
Right Boundary: $f'(x^*) \leq 0$ and $x^* = b$.

Existence: If f is continuous on $[a, b]$ then it has a minimum on $[a, b]$.

Uniqueness: If f is strictly convex on $[a, b]$ then it has a unique minimum on $[a, b]$.

Abstract Vector Space

Consider a general optimization problem:

$$\min_{x \in \mathcal{D}} \; J(x) \qquad\qquad [**]$$

where $\mathcal{D}$ is a subset of a vector space $\mathcal{V}$.

We consider functions $\zeta = \zeta(\varepsilon) : [a, b] \to \mathcal{D}$ such that the composite $J \circ \zeta$ is differentiable.

Suppose that $x^* \in \mathcal{D}$ and $J(x^*) \leq J(x)$ for all $x \in \mathcal{D}$. In addition, let ζ such that $\zeta(\varepsilon^*) = x^*$ then (*necessary conditions*):

Interior point: $\frac{\mathrm{d}}{\mathrm{d}\varepsilon} J(\zeta(\varepsilon))|_{\varepsilon=\varepsilon^*} = 0$, $\frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2} J(\zeta(\varepsilon))|_{\varepsilon=\varepsilon^*} \geq 0$, and $a < \varepsilon^* < b$.

Left Boundary: $\frac{\mathrm{d}}{\mathrm{d}\varepsilon} J(\zeta(\varepsilon))|_{\varepsilon=\varepsilon^*} \geq 0$ and $\varepsilon^* = a$.

Right Boundary: $\frac{\mathrm{d}}{\mathrm{d}\varepsilon} J(\zeta(\varepsilon))|_{\varepsilon=\varepsilon^*} \leq 0$ and $\varepsilon^* = b$.

How do we use these necessary conditions to identify “good candidates” for x^*?

Extremals and Gâteau Variations

Definitions:

Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space and let $\mathcal{D} \subseteq \mathcal{V}$.

– We say that a point $x^* \in \mathcal{D}$ is an *extramal point* for a real-valued function J on $\mathcal{D}$ if

$$J(x^*) \leq J(x) \quad \text{for all } x \in \mathcal{D} \;\vee\; J(x^*) \geq J(x) \quad \text{for all } x \in \mathcal{D}.$$

– A point $x_0 \in \mathcal{D}$ is called a *local extremal point* for J if for some $\epsilon > 0$, x_0 is an extremal point on $\mathcal{D}_\epsilon(x_0) := \{x \in \mathcal{D} \,:\, \|x - x_0\| < \epsilon\}$.

– A point $x \in \mathcal{D}$ is an *internal* (*radial*) point of $\mathcal{D}$ in the direction $v \in \mathcal{V}$ if

$$\exists \varepsilon(v) > 0 \text{ such that } \; x + \varepsilon v \in \mathcal{D} \text{ for all } |\varepsilon| < \varepsilon(v) \;\; (0 \leq \varepsilon < \varepsilon(v)).$$

– The directional derivative of order n of J at x in the direction v is denoted by

$$\delta^n J(x; v) = \frac{\mathrm{d}^n}{\mathrm{d}\varepsilon^n} J(x + \varepsilon v)|_{\varepsilon=0}.$$

– J is *Gâteau-differentiable* at x if x is an internal point in the direction v and $\delta J(x; v)$ exists for all $v \in \mathcal{V}$.

Theorem:(Necessary Conditions)

Let $(\mathcal{V}; \|\cdot\|)$ be a normed linear space. If J has a (local) extremal at a point x^* on $\mathcal{D}$ then $\delta J(x^*, v) = 0$ for all $v \in \mathcal{V}$ such that (i) x^* is an internal point in the direction v and (ii) $\delta J(x^*, v)$ exists.

This result is useful if there is "enough" directions v so that the condition $\delta J(x^*, v) = 0$ can determine x^*.

Examples:

– Find the extremal points for

$$J(y) = \int_a^b y^2(x)\,\mathrm{d}x$$

on the domain $\mathcal{D} = \{y \in C[a,b] \,:\, y(a) = \alpha$ and $y(b) = \beta\}$.

– Find the extremal for

$$J(P) = \int_a^b P(t)\,D(P(t))\,\mathrm{d}t$$

on the domain $\mathcal{D} = \{P \in C[a,b] \,:\, \dot{P}(t) \leq \xi\}$.

Extremal with Constraints

Suppose that in a normed linear space $(\mathcal{V}, \|\cdot\|)$ we want to characterize extremal points for a real-valued function J on a domain $\mathcal{D} \subseteq \mathcal{V}$. Suppose that the domain is given by the *level set* $\mathcal{D} := \{x \in \mathcal{V} \,:\, G(x) = \psi\}$, where G is a real-valued function on $\mathcal{V}$ and $\psi \in \mathbb{R}$.

Let x^* be a (local) extremal point. We will assume that both J and G are defined in a neighborhood of x^*.

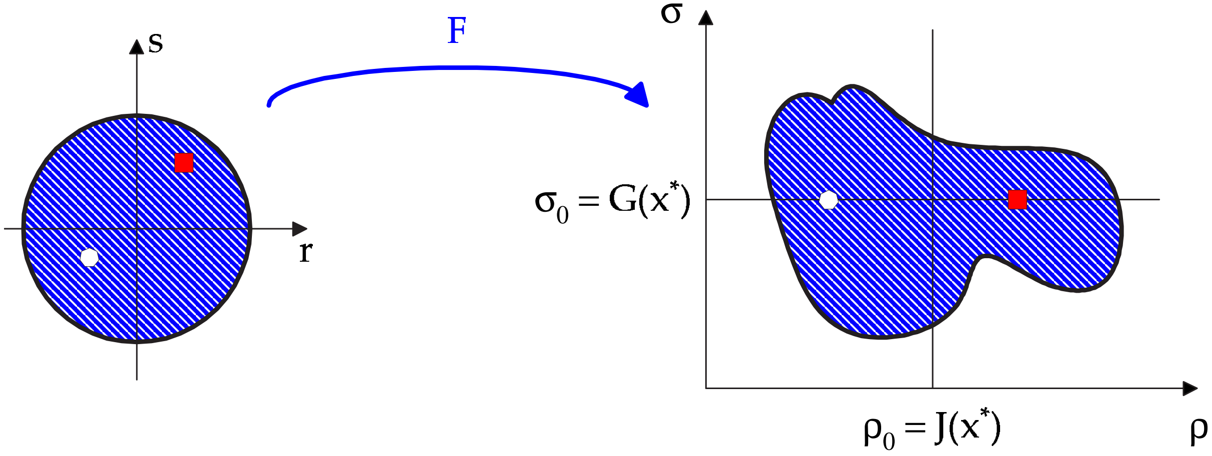
We pick an arbitrary pair of directions v, w and a define the mapping

$$F_{v,w}(r,s) := \begin{pmatrix} \rho(r,s) \\ \sigma(r,s) \end{pmatrix} = \begin{pmatrix} J(x^* + rv + sw) \\ G(x^* + rv + sw) \end{pmatrix}$$

which is well defined in a neighborhood of the origin.

Suppose F maps a neighborhood of 0 in the (r, s) plane into an neighborhood of $(\rho^*, \sigma^*) := (J(x^*), G(x^*))$ in the (ρ, σ) plane.

Then x^* cannot be an extremal point of J.



This condition is assured if F has an inverse which is continuous at (ρ^*, σ^*).

Theorem: For $\bar{x} \in \mathbb{R}^n$ and a neighborhood $\mathcal{N}(\bar{x})$, if a vector valued function $F : \mathcal{N}(\bar{x}) \rightarrow \mathbb{R}^n$ has continuous first partial derivatives in each component with nonvanishing Jacobian determinant at $\bar{x}$, then F provides a continuously invertible mapping between a neighborhood of $\bar{x}$ and a region containing a full neighborhood of $F(\bar{x})$.

In our case, $\bar{x} = 0$ and the Jacobian matrix of F is given by

$$\nabla F(0,0) = \begin{pmatrix} \delta J(x^*; v) & \delta J(x^*; w) \\ \delta G(x^*; v) & \delta G(x^*; w) \end{pmatrix}$$

Then if $|\nabla F(0,0)| \neq 0$ then x^* cannot be an extremal point for J when constraint to the level set defined by $G(x^*)$.

Definition: In a normed linear space $(\mathcal{V}, \|\cdot\|)$, the Gâteau variations $\delta J(x,v)$ of a real valued function J are said to be *weakly continuous* at $x^* \in \mathcal{V}$ if for each $v \in \mathcal{V}$ $\delta J(x;v) \longrightarrow \delta J(x^*;v)$ as $x \longrightarrow x^*$.

Theorem:(Lagrange) In a normed linear space $(\mathcal{V}, \|\cdot\|)$, if a real valued functions J and G are defined in a neighborhood of x^*, a (local) extremal point for J constrained by the level set $G(x^*)$, and have there weakly continuous Gâteau variations, then either
a) $\delta G(x^*; w) = 0$, for all $w \in \mathcal{V}$, or
b) there exists a constant $\lambda \in \mathbb{R}$ such that $\delta J(x^*, v) = \lambda \delta G(x^*; v)$, for all $v \in \mathcal{V}$.

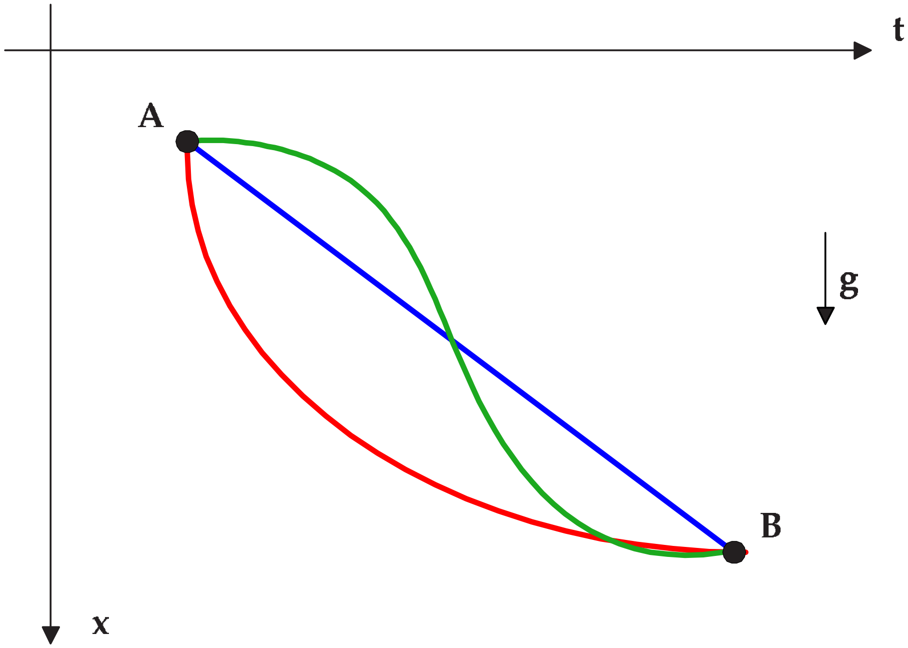
Example: Find the extremal for

$$J(P) = \int_0^T P(t)\, D(P(t))\, \mathrm{d}t$$

on the domain $\mathcal{D} = \{P \in C[0,T] \,:\, \int_0^T D(P(t))\, \mathrm{d}t = I\}$.

Classical Calculus of Variations
Historical Background

Johann Bernoulli (1696)- Brachistochrone: Find the planar curve which would provide the faster time of transit to a particle sliding down it under the action of gravity.



Solutions by Jakob Bernoulli, Newton, Euler, Leibniz, and L'Hospital.

Geodesic Problems: Find the shortest path in a given domain connecting two points of it.

The Simplest Problem in Calculus of Variations

$$J(x) = \int_a^b L(t, x(t), \dot{x}(t))\, \mathrm{d}t,$$

where $\dot{x}(t) = \frac{\mathrm{d}}{\mathrm{d}t} x(t)$. The *variational integrand* is assumed to be smooth enough (*e.g.*, at least C^2).

Examples:

– Geodesic: $L = \sqrt{1 + \dot{x}^2}$ – Brachistochrone: $L = \sqrt{\frac{1+\dot{x}^2}{x-\alpha}}$

– Minimal Surface of Revolution: $L = x\,\sqrt{1 + \dot{x}^2}$.

Admissible Solutions:

A function $x(t)$ is called *piecewise* C^n on $[a, b]$, if $x(t)$ is C^{n-1} on $[a, b]$ and $x^{(n)}(t)$ is piecewise continuous on $[a, b]$,*i.e*, continuous except on a finite number of points. We denote by $\mathcal{H}[a, b]$ the vector space of all real-valued piecewise C^1 function on $[a, b]$ and by $\mathcal{H}_e[a, b]$ the subspace of $\mathcal{H}[a, b]$ such that $x(a) = x_a$ and $x(b) = x_b$ for all $x \in \mathcal{H}_e[a, b]$.

Problem:

$$\min_{x \in \mathcal{H}_e[a,b]} J(x).$$

Admissible Variations or Test Functions:

Let $\mathcal{Y}[a,b] \subseteq \mathcal{H}[a,b]$ be the subspace of piecewise C^1 functions y such that

$$y(a) = y(b) = 0.$$

We note that for $x \in \mathcal{H}_e[a,b]$, $y \in \mathcal{Y}[a,b]$, and $\varepsilon \in \mathbb{R}$, the function $x + \varepsilon y \in \mathcal{H}_e[a,b]$.

Theorem: Let J have a minimum on $\mathcal{H}_e[a,b]$ at x^*. Then

$$L_{\dot{x}} - \int_a^t L_x \, \mathrm{d}\tau = \text{constant} \qquad \text{for all } t \in [a,b]. \tag{1}$$

A function $x^*(t)$ satisfying (1) is called *extremal*.

Corollary:(**Euler's Equation**)
Every extremal x^* satisfies the differential equation

$$L_x = \frac{\mathrm{d}}{\mathrm{d}t} L_{\dot{x}}.$$

Example: Production-Inventory Control

Consider a firm that operates according to a make-to-stock policy during a planning horizon $[0, T]$. The company faces an exogenous and deterministic demand with intensity $\lambda(t)$. Production is costly; if the firm chooses a production rate μ at time t then the instantaneous production cost rate is $c(t, \mu)$. In addition, there are holding and backordering costs. We denote by $h(t, I)$ the holding/backordering cost rate if the inventory position at time t is I. We suppose that the company starts with an initial inventory I_0 and tries to minimize total operating costs during the planning horizon of length $T > 0$ subject to the requirement that the final inventory position at time T is I_T.

a) Formulate the optimization problem as a calculus of variations problem.

b) What is Euler's equation?

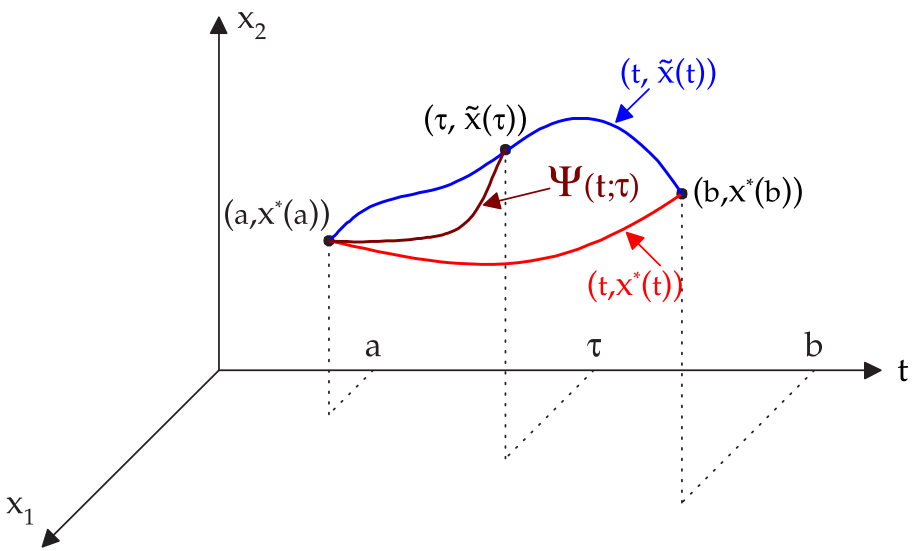
Sufficient Conditions: Weierstrass Method

Suppose that x^* is an extremal for

$$J(x) = \int_a^b f(t, x(t), \dot{x}(t))\, dt := \int_a^b f[x(t)]\, dt$$

on $\mathcal{D} = \{x \in C^1[a,b] : x(a) = x^*(a);\, x(b) = x^*(b)\}$. Let $\tilde{x}(t) \in \mathcal{D}$ be an arbitrary feasible solution.

For each $\tau \in (a,b]$ we define the function $\Psi(t;\tau)$ on (a,τ) such that $\Psi(t;\tau)$ is an extremal function for f on (a,τ) whose graph joins $(a, x^*(a))$ to $(\tau, \tilde{x}(\tau))$ and such that $\Psi(t;b) = x^*(t)$.



We define

$$\sigma(\tau) := -\int_a^\tau f[\Psi(t;\tau)]\,\mathrm{d}t - \int_\tau^b f[\tilde{x}(t)]\,\mathrm{d}t,$$

which has the following properties:

$$\sigma(a) = -\int_a^b f[\tilde{x}(t)]\,\mathrm{d}t = -J(\tilde{x}) \quad \text{and} \quad \sigma(b) = -\int_a^b f[\Psi(t,b)]\,\mathrm{d}t = -J(x^*).$$

Therefore, we have that

$$J(\tilde{x}) - J(x^*) = \sigma(b) - \sigma(a) = \int_a^b \dot{\sigma}(\tau)\,\mathrm{d}\tau,$$

so that a sufficient condition for the optimality of x^* is $\dot{\sigma}(\tau) \geq 0$. That is,

$$\begin{aligned} \dot{\sigma}(\tau) \quad &:= \quad \mathcal{E}(\tau, \tilde{x}(\tau), \dot{\Psi}(\tau;\tau), \dot{\tilde{x}}(\tau)) \\ &= \quad f[\tilde{x}(\tau)] - f(\tau, \tilde{x}(\tau), \dot{\Psi}(\tau;\tau)) - f_{\dot{x}}(\tau, \tilde{x}(\tau), \dot{\Psi}(\tau;\tau)) \cdot (\dot{\tilde{x}}(\tau) - \dot{\Psi}(\tau;\tau)) \geq 0 \end{aligned}$$

Weierstrass' formula