Minimization (maximization) with constraints. Lagrange multipliers

Example in $\mathbb{R}^2$

$I: \mathbb{R}^3 \to \mathbb{R}$,

$I(x, y, z) = 2$.

$J: \mathbb{R}^3 \to \mathbb{R}$.

$J(x, y, z) = x^2 + y^2 + z^2 - 1$.

$$\min I(x, y, z) = \begin{cases} (x_0, y_0, z_0) = (0, 0, -1) \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

We also have $3 \lambda \in \mathbb{R}$:

$$\nabla I(x_0, y_0, z_0) = \lambda \nabla J(x_0, y_0, z_0)$$

because the differential of $I$ at the minimum must be zero on the tangent space (plane)

$$DI[v] = \nabla I \cdot v = 0 \quad \forall v \in T(x_0, y_0, z_0)$$

$$A = \{ (x, y, z) \in \mathbb{R}^3 \mid J(x, y, z) = 0 \}$$

since the tangent space is defined by the level tangent to curves in $A$ passing through the point $(x_0, y_0, z_0)$.

But $T(x_0, y_0, z_0) A = \{ v \in \mathbb{R}^3 \mid \nabla J(x_0, y_0, z_0) \cdot v = 0 \}$

$$= \{ \nabla I \text{ and } \nabla J \text{ must be parallel} \}$$
Worse generally:

**Theorem (in $\mathbb{R}^n$)** Let $x_0 \in \mathbb{R}^n$ solve.

$$\min \ I(x)$$

$$x \in A = \{ x \in \mathbb{R}^n : \ T_1(x), \ldots, T_m(x) = 0 \}$$

Then if $I$ is differentiable at $x_0$, $T_1, \ldots, T_m$ are $C^1$ in a neighborhood of $x_0$ and $\nabla T_1(x_0), \ldots, \nabla T_m(x_0)$ are linearly independent then there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that

$$\nabla I(x_0) = \sum_{k=1}^{m} \lambda_k \nabla T_k(x_0)$$

Remark: $\lambda_k$ are called Lagrange multipliers.

**Proof** Construct curves $v(t) \in A$ passing through $x_0$ and suppose that $x_0$ is a minimum for $I$ along these curves. I will focus on $m = 1$, general case $m \in \mathbb{N}$.

Fix $w \in \mathbb{R}^n$ such that $\nabla T_1(x_0), w \neq 0$ (since $\nabla T_1(x_0)$ is linearly independent $\iff \nabla T_1(x_0) \neq 0$ hence such a $w$ exist).

For any $v \in \mathbb{R}^n$ consider $j : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$.

$$j(t, \theta) = J(t_0 + t_v, \theta w)$$

-2-
which by hypotheses is \( C' \) in a neighborhood of \((0,0)\).

Moreover: \( j(0,0) = J(x_0) = 0 \) since \( x_0 \in \Omega \).

\[
\frac{\partial}{\partial \theta}(0,0) = \nabla J(x_0) \cdot \mathbf{w} \neq 0
\]

By Hyphlic Function Theorem \( \exists \delta > 0 \) and \( \Theta: (-\delta, \delta) \rightarrow \mathbb{R} \) such that

\[
j(c, \Theta(c)) = J(x_0 + c \mathbf{v} + \Theta(c) \mathbf{w}) = 0 \quad \forall c \in (-\delta, \delta)
\]

Hence \( \mathbf{x} \mapsto x_0 + c \mathbf{v} + \Theta(c) \mathbf{w} \) is a curve in \( \mathcal{A} \).

\[
c \in (-\delta, \delta) \Rightarrow i(c) = \mathbf{x} \quad (x_0 + c \mathbf{v} + \Theta(c) \mathbf{w}) \text{ has a minimum}
\]
at \( c = 0 \)

\[
= \quad \Rightarrow \quad \sigma = \frac{dx}{dc}(0) = \nabla I(x_0) \cdot [\mathbf{v} + \Theta'(0) \mathbf{w}]
\]

But \( \Theta'(0) \) can be calculated. Now:

\[
\Theta = \frac{dx}{dc}(c) = \nabla J(x_0 + c \mathbf{v} + \Theta(c) \mathbf{w}) \cdot [\mathbf{v} + \Theta'(c) \mathbf{w}]
\]

\[
\Rightarrow \quad \Theta'(0) = -\frac{\nabla J(x_0) \cdot \mathbf{v}}{\nabla J(x_0) \cdot \mathbf{w}}
\]

For \( \lambda = \frac{\nabla I(x_0) \cdot \mathbf{w}}{\nabla J(x_0) \cdot \mathbf{w}} \) (a fixed real number)

we have.

\[-3-\]
\[ \nabla I(x_0) \cdot v = \lambda \nabla J(x_0) \cdot v \quad \forall v \in \mathbb{R}^n \]

\[ \Rightarrow \nabla I(x_0) = \lambda \nabla J(x_0) \quad \text{q.e.d.} \]

**Example in Banach spaces**

Consider \( I : H_0^1(\Omega) \rightarrow \mathbb{R} \), \( \Omega \in \mathbb{R}^n \) open.

\[ I(v) = \int_{\Omega} \frac{1}{2} |\nabla v|^2 - \frac{1}{p} (|v|^{p-2} v) dx \]

\[ J : H_0^1(\Omega) \rightarrow \mathbb{R} \]

\[ J(v) = \int_{\Omega} |\nabla v|^2 dx - 1 \]

**Theorem** Let \( v \) solve

\[ \min_{v \in V} I(v) \]

\( V \in U = \{ v \in H_0^1(\Omega) \mid (v) = 0 \} \).

the \( v \) is a weak soln of

\[ -\Delta v - |v|^{p-2} v = \lambda v \]

for some \( \lambda \in \mathbb{R} \).
More generally, if \( X \) is a Banach space (over \( \mathbb{R} \) or \( \mathbb{C} \)), and \( I : X \rightarrow \mathbb{R} \), \( J : X \rightarrow \mathbb{R} \) have the following properties:

1. \( \mathcal{A} = \{ x \in X \mid J(x) = 0 \} \neq \emptyset \).

2. \( x_0 \) is a local minimum for \( I \).

3. \( J \) is \( C^1 \) in a neighborhood of \( x_0 \) and \( DJ(x_0) \neq 0 \), i.e., the Fréchet differential of \( J \) at \( x_0 \) is not the zero map.

4. \( DI(x_0) \) exists, i.e., the Fréchet differential of \( I \) at \( x_0 \) exists.

Then there exists \( \lambda \in \mathbb{R} \) such that

\[
DI(x_0) = \lambda \cdot DJ(x_0)
\]

Note: For \( F : X \rightarrow Y \) between Banach spaces, \( DF(x_0) : X \rightarrow Y \) is an linear continuous map such that

\[
\lim_{x \to x_0} \frac{F(x) - F(x_0) - DF(x_0)(x - x_0)}{\|x - x_0\|} = 0.
\]
Proof. Choose \( w \in X : D J (x_0) \Sigma w \neq 0 \).

For each \( \bar{u} \in X \) let \( j : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \) be given by
\[
j(\bar{v}, \theta) = J(x_0 + \bar{v} u + \theta \Sigma w).
\]

Repeat the argument in the \( \mathbb{R}^n \) case using:
\[
3^0 = T \ j \text{ is continuous in a neighborhood of } (0, 0).
\]

We will have:
\[
\frac{d j}{d \theta} (\bar{v}, \theta) = D J(x_0 + \bar{v} u + \theta \Sigma w) \Sigma w
\]

in a neighborhood of \((0, 0)\)
\[
\frac{d j}{d \bar{v}} (\bar{v}, \theta(\bar{v})) = D J(x_0 + \bar{v} u + \theta(\bar{v}) \Sigma w) \Sigma u + \theta'(\bar{v}) \Sigma w
\]

in a neighborhood of \( \bar{v} = 0 \).
\[
\frac{d j}{d \bar{v}} (0) = \frac{d}{d \bar{v}} \left. J(x_0 + \bar{v} u + \theta(\bar{v}) \Sigma w) \right|_{\bar{v} = 0}.
\]
\[
= D I(x_0) \Sigma u + \theta'(0) \Sigma w
\]

and with the notation \( \Lambda = \frac{D I(x_0) \Sigma w}{D J(x_0) \Sigma w} \).

We get:
\[
D I(x_0) \Sigma u = \Lambda D J(x_0) \Sigma u \quad \forall \ u \in X
\]

\( \Rightarrow \) maps \( D I(x_0) \) and \( \Lambda D J(x_0) : X \to \mathbb{R} \) coincide.!!
To apply this general theory to the example one must know how to compute Frechet differentials.

\[
\text{Proportion } \lim_{\Delta v \to 0} \frac{F(x_0 + \Delta v) - F(x_0)}{\Delta v} = dF(x_0)[tv] = dF(x_0)[v].
\]

\(dF(x_0)[tv]\) defined above is called the Gateaux differentiable if it exist in any direction \(tv \in X\) (it plays the role of directional derivatives in \(\mathbb{R}^n\)).

Definition \(B(X,Y) = \{L: X \to Y | L \text{ linear and continuous} \}\).

Theorem \(B\) is a Banach space with norm \(\|L\| = \sup \|L(x)\| \text{ if } \|x\| \leq 1\).

Theorem If \(x_0 \in V\) such that 
\(dF(x) \in B(x, y) \forall x \in V\),
\(dF: V \to B(x, y)\) is continuous at \(x_0\), then the Frechet differential of \(F\) at \(x_0\) exits and \(DF(x_0) = dF(x_0)\).
Def: If $\mathbf{DF}: V \rightarrow \mathbb{B}(x,y) \Rightarrow \mathbf{dF}: V \rightarrow \mathbb{B}(x,y)$ is continuous then $F$ is of class $C^1$.

Back to the example

$$
\lim_{\varepsilon \rightarrow 0} \frac{I(U + \varepsilon v) - I(U)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \varepsilon
$$

$$
= \lim_{\varepsilon \rightarrow 0} \int_U \frac{1}{2} \left( (\nabla U + \varepsilon \nabla v) \cdot (\nabla U + \varepsilon \nabla v) \right)
- \frac{1}{p} \left( (U + \varepsilon v)^2 \right)^{\frac{p}{2}} \, dx.
$$

Dominated

$$
= \int_U \nabla U \cdot \nabla v - \|U\|^{p-2} U v \, dx.
$$

Concern these.

For each $U \in H_0^1(U)$ the map

$$
H_0^1(U) \ni U \longmapsto \int_U \nabla U \cdot \nabla v - \|U\|^{p-2} U v \, dx
$$

is linear and continuous.

Moreover this map depends continuously on $U$.

\[2) \text{ I is } C^1 \text{ on } H_0^1(U) \text{ and } \text{ continuous.}\]
\( \partial I (u) \left[ v \right] = \int \nabla u \cdot \nabla v - 1 u^{1-p-2} u \cdot v \, dx. \)

Similarly, \( J \) is \( C^1 \) on \( H^1_0(\Omega) \) and:

\[ \partial J (u) \left[ v \right] = 2 \int \nabla u \cdot \nabla v \, dx. \]

By the Theorem if \( u \) solves

\[ \min \ I (u), \quad J(u) = 0, \quad \|u\|_{L^2} = 1 \]

then \( \exists \lambda \in \mathbb{R} : \)

\[ \partial I (u) \left[ \lambda v \right] = \frac{\lambda}{2} \partial J (u) \]

\( \Rightarrow \)

\[ \int \nabla u \cdot \nabla v - (1 u^{1-p-2} u + \lambda u) v \, dx = 0 \quad \forall v \in H^1_0(\Omega) \]

\( \Rightarrow \) \( u \) is a weak sol of \( -\Delta u - 1 u^{1-p-2} u = \lambda u \)
Solving the minimization problem: \( \Omega \leq 10^{-24} \text{open} \)

\[
\min_{v \in H_0^1(\Omega)} I(v) = \int_{\Omega} \left( \frac{1}{2} |Dv|^2 - \frac{1}{p} |10v|^p \right) dx
\]

\[\|v\|_{L^2} = 1\]

Uses Gagliardo-Nirenberg inequality, namely

\[
\|v\|_{L^p} \leq c \|v\|_{L^2} \|v\|_{L^2}
\]

and require \( p < \frac{2n+4}{n} \) i.e. \( p-2 < \frac{4}{n} \) (Sobolev inequality). For lower bounds for \( I(v) \) and coercivity

On \( \Omega \) bounded, minimizing sequences \( \{v_m\} \subset H_0^1(\Omega) \)

are bounded and converge in \( L^p, 2 \leq p < \frac{2n}{n-2} \), on subsequence.

due to the compact embedding

\[
H_0^1(\Omega) \subset \text{comp} \ L^2(\Omega), \quad 2 \leq p < \frac{2n}{n-2}
\]

hence the existence of the minimizer follows from

Same argument as for Lagrange (see text book)

Even though the conditions:

\[
L(1,2) \leq C(10^{1/2} + 10^{1/2})
\]

\[
L(1,2) \geq 2 \left| 10^{1/2} - 1 \right|
\]

are not satisfied (\( 2 \) must be 2 from the 2nd, but \( 1/2 > 2 \) violates the 1st impossible.
On unbounded domains $D$ the term $v^p$ in $I$ can no longer be treated as in the textbook. It requires concentration compactness to obtain a minimizer.

See

math.iiisci.ernet.in/teaching/math555

Lectures 13 and 14 on the course advertised on the next page for details.
Math 505 MNA Spring 2019

Methods in Nonlinear Analysis and Applications to Differential Equations

Class time: TR
Lecturer: Eduardo Kirr, e-mail: ekirr@illinois.edu

Description: The first part of the course will focus on fixed point theorems based on degree theory and their applications to differential equations. The (Brouwer) degree for maps between finite dimensional spaces has deep roots in homotopy and homology and you might have encountered it when studying these theories. While I will briefly discuss how it is introduced there, I will present a more axiomatic view of degree theory which can be generalized to (nonlinear) maps between infinite Banach spaces having certain compactness properties (completely continuous, properly bounded, etc.). I will discuss its powerful consequences: Brouwer, Borsuk-Ulam, Ham and Sandwich Theorems in finite dimensions and the Leray-Schauder type fixed point theorems in infinite dimensions. Their applications to modern local and global bifurcation theories and to solutions of differential equations will be emphasized. This part of the course will culminate with the contribution made by the degree theory in understanding the collapse of Tacoma-Narrows Bridge, a phenomenon that was not predicted by laboratory simulations or the structural and dynamical stability theory preceding its construction.

The second part of the course will focus on contraction principle and variational methods. While independent, this part will extend some of the results in the first part to non-completely continuous (non-compact) maps. Metric and Banach spaces will be reviewed and Calculus in Banach spaces will be introduced before proving the Banach fixed point theorem and its consequences e.g., the implicit function theorem (IFT) in (infinite) dimensional Banach spaces. The application of the contraction principle to existence, uniqueness, continuous dependence of data and stability for solutions of evolution equations (including systems of ordinary and partial differential equations) will be briefly discussed while the applications of the IFT to Lyapunov-Schmidt decomposition and local bifurcation theory will be presented in detail with examples from nonlinear optics and statistical physics. Extensions of the local results to non-perturbative regimes via the global bifurcation theory for real analytical maps will also be discussed. The variational methods will be based on the rigorous calculus in Banach spaces which replaces the more common but ad-hoc “calculus of variation” in which the definition of “variation” seems to change from problem to problem. Moreover, when it comes to applications in finding certain equilibria or periodic solutions in partial differential equations, the variational methods have to cope with non-convex functional and non-compact constraints. We will discuss how to compensate for these shortcomings via Rellich or concentration compactness and then apply the classical theory which of course will be introduced.

References: I will follow my own notes (posted online) based on the following references:

1. Topics in Nonlinear Functional Analysis by L. Nirenberg


Grading Policy: There are no homework assignments or exams for this course. The participants will be asked to make a presentation on the applications of these techniques to a nonlinear differential equation, preferably from their own research area. Grades will be based on class activity, and on the quality of the presentation.