Proposal Description

RUI: Variational Analysis for Optimizations on Metric Spaces and Applications

Updated on Feb. 17, 2007

Libin Mou
Department of Mathematics, Bradley University, Peoria, IL 61625

Sections

Page

Section 1. Problem Statement and Objectives 1-2
Section 2. Related Previous Works 3-5
Section 3. Proposed Research Plans 5-14
Section 4. Broader Impacts of the Proposal 15
Section 5. Cited References 16-19

Section 1. Problem Statement and Objectives

An optimization problem on a metric space can be stated as follows.

**Problem.**

\[
\text{minimize } J(w), \ w \in \mathcal{W} \ \text{subject to } S(w) \in Q,
\]

where \((\mathcal{W}, d)\) is a complete metric space, \((Z, \| \cdot \|)\) a Banach space, \(J(\cdot) : \mathcal{W} \rightarrow \mathbb{R}\) the objective function, \(S(\cdot) : \mathcal{W} \rightarrow Z\) the constraint map, and \(Q \subset Z\) a closed subset. Many optimization problems (in calculus of variations, optimal controls, game theories, mathematical programming, etc.) with different types of constraints can be formulated like Problem (1); as an example, see the stochastic optimal controls in Section 3.3.

Existing variational analyses (linear and nonlinear, smooth and nonsmooth) are exclusively established on *Banach spaces* (or more specially, Hilbert and Euclidean spaces). Such analyses can be *indirectly* applied to certain optimization problems on metric spaces (e.g. optimal controls), however, tedious "bookkeeping" is often required to convert the problems into the forms to which these analyses apply; see Section 3.2 in the cited reference [8] for an example.

Note that there are geometric analyses developed on metric spaces (see [1] and references therein), but they are not appropriate for solving optimization problems.
Can a variational analysis be developed for solving optimization problems on metric spaces? Are there appropriate notions of derivatives/differentials on metric spaces for such an analysis?

In my recent work [27] (with M. McAsey), notions of sequential derivative and sequential strict derivative were defined for maps on metric spaces, and a Lagrange multiplier rule for Problem (1) was proved in the case that $Z$ has a strictly convex dual $Z^*$ and $Q$ is convex and finitely codimensional. The multiplier rule in [27] guarantees the existence of multiplier(s) which satisfy certain inequalities involving all sequential strict derivatives. As an application of the multiplier rule, a unified proof of the maximum principle is given for deterministic optimal controls with general constraints (isoperimetric and pointwise) under the strict differentiability of the involved data. The strict differentiability condition is by far the weakest assumption for a classical maximum principle of optimal control with pointwise constraints. (In special cases, the assumptions can be further weakened; see [3] and [17].)

I will develop a variational analysis for solving optimization problems in terms of sequential derivatives and subdifferentials on metric spaces and apply the analysis to stochastic optimal controls with general constraints to prove a maximum principle. Specifically, the following plans are proposed.

I. Short-term plans (2007-2009)
   (1) Thoroughly examine properties of sequential derivative and sequential strict derivative for maps on metric spaces. In particular, I will verify that these notions of derivatives are appropriate for developing a variational analysis on metric spaces. In addition, basic rules of sequential differentiation will be established. The notion of subdifferential on metric spaces will be defined in Long-term plan in Section 3.4.
   (2) Generalize the multiplier rule (Theorem 1 Section 2.1) to general constraint sets $Q$'s; this will significantly expand the scope of applications of multiplier rules.
   (3) Apply the multiplier rules (Theorem 1 and its generalizations) to stochastic optimal controls with general constraints.

II. Long-term plans (2008 and later)
   (4) Subdifferentials and variational analysis on metric spaces
   (5) Multi-objective optimizations on metric spaces
   (6) Further applications
In summary, this research is to develop a variational analysis specifically for solving optimization problems on metric spaces. Such an analysis bears direct applications to optimization problems, especially optimal controls with various constraints. The research complements existing generalized geometric analyses on metric spaces, and extends (to some degree) existing variational analyses on Banach spaces.

Section 2. Related Previous Works

(The investigator's papers are available at http://hilltop.bradley.edu/~mou/moupaper.html)

(This proposal with updates is available at http://hilltop.bradley.edu/~mou/proposal2006.html)

2.1. A Multiplier rule for optimization on metric spaces [27]

Let \( F : (\mathcal{W}, d) \to (X, || \cdot ||) \) be a map from a complete metric space to a Banach space. The following notions of sequential derivative, \( \delta \)-sequential derivative and sequential strict derivative were introduced in [27]. Denote \( \mathbb{N} = \{1, 2, \cdots\} \). For \( w \in \mathcal{W} \), let \( \mathcal{S}_w \) be the set of sequences \( (w^i, d^i)_{i \in \mathbb{N}} \) with \( w^i \in \mathcal{W}, d^i \in (0, \infty) \), and \( d(w^i, w) \leq d^i \downarrow 0 \) as \( i \to \infty \).

Definition 1. An element \( x \in X \) is called a sequential derivative of \( F \) at \( w \in \mathcal{W} \) if

\[
x = \lim_{i \to \infty} \frac{F(w^i) - F(w)}{d^i}
\]

for some \( (w^i, d^i)_{i \in \mathbb{N}} \in \mathcal{S}_w \). The sequence \( w^i \to w \) defines a "direction" in \( \mathcal{W} \). The set \( DF(w) \) of all sequential derivatives at \( w \) can be considered as a generalization of directional derivatives of maps on Banach spaces; see Proposition 1 in Section 3.

Definition 2. For a given \( \delta \geq 0 \), \( x \in X \) is called a sequential \( \delta \)-derivative of \( F \) at \( w \) if

\[
\limsup_{i \to \infty} || \frac{F(w^i) - F(w)}{d^i} - x || \leq \delta.
\]

(2)

for some \( (w^i, d^i)_{i \in \mathbb{N}} \in \mathcal{S}_w \). The set of all \( \delta \)-derivatives of \( F \) at \( w \) is denoted as \( D^\delta F(w) \).

Definition 3. Now fix a \( w_0 \in \mathcal{W} \). If for every \( w \in \mathcal{W} \), \( x \) is a sequential \( \delta(w) \)-derivative of \( F \) at \( w \) for some \( \delta(w) \), for which \( \delta(w) \to 0 \) as \( d(w, w_0) \to 0 \), then \( x \) is called a sequential strict derivative of \( F \) at \( w_0 \). The set of all sequential strict derivatives at \( w_0 \) is denoted by \( D_s F(w_0) \).
In Definitions 1-3 above we may let \( d_i = d(w^i, w) \), however, the flexibility of choosing \( d_i \geq d(w^i, w) \) is a big convenience in applications; please see the remark at the end of [27]. These notions of derivatives have the desired properties for studying optimization problems (with or without constraints); see Section 3.1 for more details. The following multiplier rule was proved in [27] for Problem (1) above. (The concepts of strict convexity and finite codimensionality are defined, for example, in [23, pp. 41 and 135].)

**Theorem 1. (Multiplier Rule).** Suppose that \( w_0 \) is a minimum point of \( J(\cdot) \) subject to \( S(\cdot) \in Q \). Suppose that \( Z \) has strictly convex dual \( Z^* \) and \( Q \subset Z \) is closed, convex and finitely codimensional. Then there exist multipliers \( (\psi^0, \psi) \in [0, \infty) \times Z^* \) such that

\[
\begin{align*}
|\psi^0|^2 + \|\psi\|^2_{Z^*} & > 0, \\
\psi^0 z^0 + \langle \psi, z \rangle & \geq 0 \text{ for all } (z^0, z) \in D_s(J, S)(w_0), \\
\langle \psi, \eta - S(w_0) \rangle & \leq 0 \text{ for all } \eta \in Q,
\end{align*}
\]

(3.1) (3.2) (3.3)

where \( \langle \psi, z \rangle \) is the pairing between elements of \( Z^* \) and \( Z \).

Here (3.1) says that \( (\psi^0, \psi) \) are nontrivial, (3.2) is the inequality satisfied by the multipliers and strict derivatives, and (3.3) is the so-called transversality condition, which can be used to help determine \( \psi \). Theorem 1 is the first multiplier rule for an optimization problem on a metric space, although there have been works on abstract multiplier rules; see [15] and [26].

As a special application of Theorem 1, let \( X, Z \) be Banach spaces, \( W \subset X \) a closed subset, \( F : X \to \mathbb{R} \) the objective function, \( S : X \to Z \) a map, and \( Q \subset Z \) a closed and convex subset. Consider the following optimization problem:

**Problem (4)**

\[ \text{minimize } F(x) \text{ subject to } x \in W \text{ (a geometric constraint)} \]
\[ \text{and } S(x) \in Q \subset Z \text{ (operator constraint)} \]

Problem (4) is an example of Problem (1) with \( J = F|_W \), which has been extensively studied; please see [33, Chapter 5] (or [6], [8], [9], [42]) for various types of results and extensive references. An application of Theorem 1 to Problem (4) leads to a necessary condition for minimum points of (4) in terms of the strict derivative \( D_s(F|_W) \). It would be desirable to represent \( D_s(F|_W) \) in terms of \( D_sF \) and the geometry of \( W \). (This is proposed in Section 3.4.)

Applying Theorem 1 is fairly convenient. In fact, one needs only to "calculate" the set \( D_s(J, S)(w_0) \) of strict derivatives by constructing appropriate variations of elements \( w \in W \). Then for each \( (z^0, z) \in D_s(J, S)(w_0) \), (3.2) gives an inequality satisfied by the multipliers \( (\psi^0, \psi) \). For an optimal control problem, (3.2) can be represented in terms of the Hamiltonian combined with the multipliers, which gives the so-called maximum principle.
My *other* related previous works are listed below with very brief descriptions.

2.2. Two-person zero-sum linear quadratic stochastic differential games by a Hilbert space method [34]. This will be related to the long-term plan proposed in Section 3.5.

2.3. A unified proof of classical maximum principle for deterministic optimal controls with general state constraints [27]. By applying the multiplier rule (Theorem 1) a unified proof was given for the classical maximum principle for deterministic optimal controls with general state constraints, assuming the strict differentiability of involved data.

2.4. Variational formula for stochastic controls and some applications [35]. In this work, a variational formula was proved for stochastic controls. As an application, known maximum principles for stochastic optimal controls in [38] [49] were generalized to stochastic controls with more general controlled equations and objective functionals. In addition, extremum principles were derived for zero-sum and non-zero-sum stochastic games. In Section 3.3, I will propose an application of the multiplier rules to stochastic optimal controls with *pointwise* constraints, which are usually much harder to handle than isoperimetric constraints.

2.5. Differential Riccati equations in stochastic quadratic differential games [28]. This paper studied the rational matrix differential equations (or generalized Riccati equations) arising from stochastic quadratic differential games. This type of equations is the most general Riccati equation ever considered in the literature. We established comparison, monotonicity, existence theorems and other fundamental properties of the solutions. As an application, we obtained the existence of periodic (including constant) solutions of the Riccati equation. This also relates to the plan in Section 3.5.

**Section 3. Proposed Research Plans**

The proposed research plans are described in this section of three parts.


3.1. Basic properties of sequential derivatives
3.2. General multiplier rules on metric spaces
3.3. Application to stochastic optimal controls with general constraints

**Part II: Long-Term Plans (2008 and later)**

3.4. Subdifferentials and variational analysis on metric spaces
3.5. Multi-objective optimizations on metric spaces
3.6. Further applications

**Part III: How the Plans Will Be Carried Out.**

This part contains the research plans to be carried out within next two years.

3.1. Basic properties of sequential derivatives

The main goal of this research is to develop a variational analysis on metric spaces using the sequential derivatives defined in Section 2 and subdifferentials to be defined in Section 3.4. The definition of subdifferentials requires a notion of dual spaces for metric spaces; see Section 3.4. To confirm that sequential derivatives are the right notions for developing the variational analysis, a thorough study of sequential derivatives is necessary. I will investigate the following properties of sequential derivatives.

3.1.1. *Fermat's theorem holds for sequential derivatives*, that is, if \( w \) is a local minimum point of a function \( F \) on \( W \), then \( DF(w) \subset [0, \infty) \). This basic property has nontrivial applications. For example, consider the case when \( W \) is a closed subset of a Banach space and \( F : X \to \mathbb{R} \) is a function, then the minimization problem with a geometric constraint:

\[
\text{minimize } F(x) \text{ subject to } x \in W \subset X
\]

is equivalent to: minimize \( F|_W \). Therefore, the Fermat's Theorem gives the necessary condition: \( DF|_W(w) \subset [0, \infty) \) for \( w \in W \) to be minimum point of (5). Similar to Problem (4), Problem (5) has been extensively studied in mathematical programming; see [33, Section 5.1.1] for a systematic account of necessary conditions for (5).

3.1.2. Sequential derivatives generalize classical derivatives on Banach spaces.

**Proposition 1.** Let \( W \) and \( X \) be Banach spaces and \( F : W \to X \). Then for every \( w, v \in W \) with \( \|v\| \leq 1 \), if the directional derivative \( F'(w; v) = \lim_{t \to 0} \frac{F(w+tv) - F(w)}{t} \) exists, then \( F'(w; v) \) is a sequential derivative, i.e., \( F'(w; v) \in DF(w) \). In addition, if \( W \) is finite dimensional and the Gateaux derivative of \( F \) at \( w \) exists, then \( DF(w) = \{ F'(w; v) : v \in W \text{ with } \|v\| \leq 1 \} \).

Similarly, sequential strict derivatives generalize classical strict derivatives on Banach spaces, as shown in [27]. So sequential derivatives are natural generalizations of classical derivatives on metric spaces. Here are two illustrative examples.

**Example** (i) Let \( G(x) = |x|\sin^2 \frac{1}{x} \) for \( x \in \mathbb{R} \), \( G(0) = 0 \), then \( DG(0) = [0, 1] \) and \( D_sG(0) = \{0\} \). Note that Fermat's theorem holds at the minimum point \( x = 0 \) of \( G \).

**Example** (ii) Let \( a, 0 \neq b \in \mathbb{R} \). Define \( H(x) = \begin{cases} \frac{ax}{b+ax} & \text{if } x \text{ is rational} \\ \frac{ax}{b+ax} & \text{if } x \text{ is irrational} \end{cases} \). Then \( H(x) \) is nowhere continuous, but \( DH(x) = D_sH(x) = [-|a|, |a|] \). This follows from the fact that the rational numbers and irrational numbers are both dense in \( \mathbb{R} \).
3.1.3. **Sequential derivatives are convenient to calculate.** For differentiable maps on Banach space, every directional derivative is a sequential derivative (Proposition 1 above). In general, a sequential derivative of $F$ at $w \in \mathcal{W}$ is found by constructing an appropriate sequence $(w^i, d^i)_{i \in \mathbb{N}} \in S_w$ (i.e., a variation of $w$). For example, in proofs of maximum principles for optimal controls, "spike" variations are commonly used to determine sequential derivatives of the objective functionals.

3.1.4. **Multiplier rules hold for sequential strict derivatives.** When the constraint set $Q$ is convex and finitely codimensional and $Z^+$ is strictly convex, a multiplier rule (Theorem 1 in Section 2) was proved in [27]. This is sufficient for many applications. In Section 3.2 below, I will propose generalizations of this multiplier rule.

3.1.5. **The set $DS(w)$ of sequential derivatives is essentially independent of the metrics on $\mathcal{W}$.** Indeed, if $d_1$ and $d_2$ are equivalent metrics, that is,

$$kd_1(u, v) \leq d_2(u, v) \leq Kd_1(u, v)$$

for all $u, v \in \mathcal{W}$ and some constants $0 < k \leq K < \infty$, then

$$k \ DF(d_2; w) \subset DF(d_1; w) \subset K \ DF(d_2; w),$$

where $DF(d_i; w)$ is the set of sequential derivatives of $F$ with respect to metrics $d_i$ ($i = 1, 2$), and $k \ DF(d_2; w) = \{kx : x \in DF(d_2; w)\}$. In particular, since (6) holds for $d_1 = d_2$ and all $0 < k \leq 1 \leq K < \infty$, so does (13) for $d_1 = d_2$ and all such $k$ and $K$. In other words, the set $DS(d_1; w)$ is a "truncated" cone. $D^6F(w)$ and $D_\gamma F(w_0)$ possess similar properties.

Property (3.1.5) conforms the fact that problem (1) and its minimum points should be independent of the metrics on $\mathcal{W}$. Metrics on $\mathcal{W}$ and the notions of derivatives for the maps $J$ and $S$ in (1) are just tools for determining the minimum points. Any equivalent metric on $\mathcal{W}$ can be used for developing and applying the variational analysis on $\mathcal{W}$. The paper [43] discussed various interesting notions of "metrics" on a space.

3.1.6. **Sequential derivatives obey basic calculus rules.** For example,

(i) (Sum Rule) If $F, G : \mathcal{W} \to X$ and $(x, y) \in D(F, G)(w)$, then $x + y \in D(F+G)(w)$.

(ii) (Chain Rule) Suppose $F : \mathcal{W} \to X$ and $G : X \to Y$ where $X, Y$ are Banach spaces. If $x \in DF(w)$, $G$ is Lipschitz near $F(w)$, and the directional derivative $G'(F(w); x)$ exists, then $G'(F(w); x) \in D(G \circ F(w))$, where $G \circ F : \mathcal{W} \to Y$ is the composite of $F$ and $G$.

Similar rules are expected for $\delta$-derivatives and strict derivatives.

Properties (3.1.1)-(3.1.6) demonstrate that sequential derivatives are the right notions of derivatives for developing a variational analysis on metric spaces. Some of (3.1.1)-(3.1.6) are easy to verify or have been proved in [27], while others need to be verified or generalized.
3.2. General multiplier rules on metric spaces

In the multiplier rule Theorem 1 (Section 2), \( Z^* \) is strictly convex and \( Q \) is convex and finitely codimensional. The assumption that \( Z^* \) is strictly convex is not so restrictive because every \textit{separable} Banach space can be renormed so that the dual \( Z^* \) is strictly convex; see [23, Chapter 2, Theorem 2.18] for example. Nevertheless, there are important optimizations problems with constraints like \( S(w) \in Q \), where \( Q(\subset Z) \) is nonconvex and \( Z \) is non-separable. A typical example is the end value constraint \((x(a), x(b)) \in Q \) with a non-convex subset \( Q \subset \mathbb{R}^{2n} \) in an optimal control, where \( x(\cdot): [a, b] \to \mathbb{R}^n \) is the state function.

For Problem (1) with \( \mathcal{W} \) being a Banach space, many necessary conditions (including Lagrange multiplier rules) have been obtained in terms of the normals of \( Q \); see [33, Section 5.1.2/5.1.3] and [8], [9]. A typical idea of proof is to convert the operator constraint \( S(w) \in Q \) into the geometric constraint \( w \in S^{-1}(Q) \) and use the relationship between the normals of \( S^{-1}(Q) \) and the normals of \( Q \); see [32, Corollary 1.15, Theorem 1.17]. This idea does NOT work for Problem (1) since \( \mathcal{W} \) is a metric space.

A careful examination on the proof of the multiplier rule in [27] reveals \textit{four essential} ingredients:

1. The subdifferential \( \partial d_Q(z) \) is nonempty at all \( z \) in \( Z \setminus Q \) that are close to \( S(w_0) \).
2. The subdifferential \( \partial d_Q(z) \) has a locally closed graph in the norm \( \times \) weak* topology of \( Z \times Z^* \).
3. The distance function \( d_Q(z) \) satisfies certain mean value theorem/inequality.
4. If \( S(w_0) \leftarrow z_k \in Z \setminus Q \) and \( z^* \leftarrow z^*_k \in \partial d_Q(z_k) \), then \( z^* \neq 0 \).

Now note that if \( Z \) is a \textit{weakly compactly generated} (WEG) Asplund space, then properties (1)-(3) are known for \( d_Q \) (as is a Lipschitz function on \( Z \)); see [32, Sections 2.2/3.2] for details (definitions and theorems). Property (4) holds if \( Q \) is \textit{sequentially normally compact} (SNC) or if \( Q \) is \textit{convex} and \textit{finitely codimensional} and \( Z^* \) is strictly convex (as in Theorem 1); see [32, Section 1.14] and [27]. Therefore, I make the following conjecture.

A General multiplier rule (Conjecture). Let \( w_0 \) be a minimum point of \( J(\cdot) \) on \( \mathcal{W} \) subject to \( S(\cdot) \in Q \), where \( \mathcal{W} \) is a complete metric space, \( Z \) is a weakly compactly generated (WCG) Asplund space, \( Q \subset Z \) a closed subset that is sequentially normally compact (SNC) at \( S(w_0) \) (in the case \( S(w_0) \in \partial Q \)), and \((J, S): \mathcal{W} \to \mathbb{R} \times Z \) are continuous. Then there exist
\((\psi^0, \psi) \in \mathbb{R}^+ \times Z^*\) such that \((\psi^0, \psi) \neq (0,0)\) and

\[
\begin{align*}
\psi^0 z^0 + \langle \psi, z \rangle &\geq 0 \text{ for all } (z^0, z) \in D_s(J, S)(w_0), \\
\psi &\in N(S(w_0); Q) \text{ (normal of } Q \text{ at } S(w_0))
\end{align*}
\]  \hspace{1cm} (8)

As the first step of generalizing the multiplier rule, I will prove that this conjecture is true. After this, I will further generalize the multiplier rule to Asplund generated spaces \(Z\), which form a much larger class than Asplund space. This should be possible because Asplund generated spaces retain many basic properties of subdifferentials possessed by Asplund spaces, as shown by the initial work [12]. However, the ingredients (1)-(4) above are not known yet and need to be verified one by one.

On the other hand, Theorem 1 shows that \(Z\) be WEG Asplund is NOT necessary for multiplier rule to hold if \(Q\) is convex and finitely codimensional (or SNC). Therefore, a general multiplier rule might be stated with the ingredients (1)-(4) as the only conditions, which are in terms of the distance function \(d_Q\) (not on \(Z\) or \(Q\), separately). There are a lot of research on distance functions, such as [4] [11] [14] [30] [31] [32] [39] [48], which contain some of the needed properties. Additional study on the behaviors of \(d_Q\) on \(Z\setminus Q\) is needed in order to verify the ingredients (1)-(4) under appropriate conditions. Of course, I need to confirm that these ingredients are indeed sufficient for the multiplier rule to hold in general.

### 3.3. Application to stochastic optimal controls with general constraints

There is an extensive literature on maximum principles for optimal controls; see [7], [16], [18], [19], [47], [49] for surveys. In this proposal, I consider the application of multiplier rules to stochastic optimal controls with general constraints. Let \((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})\) be a given filtered probability space, on which a \(d\)-dimensional standard Brownian motion \(W(t) \in \mathbb{R}^d\) is defined with \(\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}\) being its natural filtration, augmented by the \(\mathbb{P}\)-null sets in \(\mathcal{F}\). Let \(\mathcal{M}([0, T]; U)\) be the set of all \(\mathbb{F}\)-adapted control processes \(u(\cdot)\) on \([0, T]\) with values in some set \(U \subset \mathbb{R}^m\). Given \((x_0, u(\cdot)) \in \mathcal{W} \equiv \mathbb{R}^n \times \mathcal{M}([0, T]; U)\), consider a controlled stochastic differential equation (SDE):

\[
dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), t > 0; \quad x(0) = x_0,
\]  \hspace{1cm} (9)

and objective functionals

\[
J^i(x_0, u(\cdot)) = \mathbb{E} \left\{ \int_0^T f^i(t, x(t), u(t))dt + h^i(x_0, x(T)) \right\}, i = 0, \ldots, N; \hspace{1cm} (10)
\]

\[
J^i(x_0, u(\cdot); s) = \mathbb{E} \left\{ \int_s^T f^i(t, x(t), u(t))dt + h^i(s, x_0, x(s), x(T)) \right\}, i = N + 1, \ldots, M,\hspace{1cm}
\]
where $0 \leq N \leq M$, $0 \leq s \leq T$, $b$, $\sigma$, $f^i$ and $h^i$ are given maps with values in $\mathbb{R}^n$, $\mathbb{R}^{n \times d}$, $\mathbb{R}$ and $\mathbb{R}$, respectively, $x(\cdot)$ is the state process in $\mathbb{R}^n$, and $\mathbb{E}\{\cdot\}$ is the expectation. Note that $J^i(x_0, u(\cdot); \cdot)$ ($i = N + 1, \ldots, M$) are maps from $\mathcal{W}$ to the space $C([0, T])$ of continuous functions on $[0, T]$, which define the pointwise constraints. The isoperimetric constraints are defined by $J^i(x_0, u(\cdot)), i = 1, \ldots, N$, and $J^0(x_0, u(\cdot))$ is the objective functional to be minimized. Since maximum principles are well-known for stochastic controls with only (or with no) isoperimetric constraints ([38], [40], and [49]), I will assume that $N = 0$ (i.e., no isoperimetric constraint). In addition, $d = 1$ (scalar Brownian motion) is assumed below.

I will derive a maximum principle for a control pair $(x_0, u(\cdot))$ that minimizes the objective functional $J^0(x_0, u(\cdot))$ subject to the following constraint:

$$S(x_0, u) \equiv (J^1(x_0, u; \cdot), \ldots, J^M(x_0, u; \cdot)) \in Q_P, \quad (11)$$

where $Q_P \subset (C[0, T])^M$ is a closed set. Two specific examples of constraints (11) are

1. Let $C \subset \mathbb{R}^{2n}$ be a given closed and convex subset. Then the initial and terminal state constraint $(x(0), x(T)) \in C$ is the special case with $(h^1, \ldots, h^{2n}) = (x(0), x(T))$ and $Q = C$.

2. $Q_P = \prod_{i=1}^M [c_i(\cdot), d_i(\cdot)]$, the set of all $(y^1, \ldots, y^M) \in (C[0, T])^M$ such that for $t \in [0, T], c_i(t) \leq y^i(t) \leq d_i(t)$ for $i = 1, \ldots, M$, where $c_i(\cdot), d_i(\cdot) \in C[0, T]$ are given.

The following statement of the expected maximum principle is lengthy but similar to that of the deterministic optimal control in [27].

**Expected Stochastic Maximum Principle.** Suppose that $Q_P \subset (C[0, T])^M$ is a closed, convex and finitely codimensional subset. Let $(x_0, u(\cdot))$ be a control that minimizes $J^0(x_0, u(\cdot))$ subject to $S(x_0, u(\cdot)) \in Q_P$. Then the quantities in (1)-(3) satisfy the conditions (A)-(B) below.

1. There exist multipliers:

$$\left(\lambda^0, \Psi^1(\cdot), \ldots, \Psi^M(\cdot)\right) \in \mathbb{R}^+ \times BV([0, T], \mathbb{R}^M),$$

where $BV([0, T], \mathbb{R}^M)$ is space of functions of bounded variation.

2. There exist costates $p$, $q$ and $r$ of first order:

$$p(\cdot), q(\cdot) \in L^2_\mathcal{F}([0, T], \mathbb{R}^n) \text{ and } r(\cdot) \in BV_\mathcal{F}([0, T], \mathbb{R}^n),$$

where the subscript $\mathcal{F}$ indicates the elements in the space are $\mathcal{F}$-adapted stochastic processes.

3. There exist costates $P$, $Q$ and $R$ of second order:

$$P(\cdot), Q(\cdot) \in L^2_\mathcal{F}([0, T], \mathbb{S}^n) \text{ and } R(\cdot) \in BV_\mathcal{F}([0, T], \mathbb{S}^n)$$

where $\mathbb{S}^n$ is the set of symmetric $n \times n$ real matrices.
The optimal conditions are

\[
\begin{aligned}
(A) \quad & \left\{ \begin{array}{l}
(\lambda^0, \Psi^1(\cdot), \ldots, \Psi^M(\cdot)) \text{ is nontrivial,}
\sum_{i=1}^M \mathbb{E} \int_{[0,T]} (\xi^i(t) - J^i(\bar{x}_0, \bar{u}(\cdot); t))d\Psi^i(t) \leq 0
\end{array} \right.
\end{aligned}
\]  

for all \((\xi^1(\cdot), \ldots, \xi^M(\cdot)) \in Q_P\), where the integral is a Lebesgue-Steljes integral.

\[
\begin{aligned}
(B) \quad & H(t, \bar{x}(t), v, p(t) + r(t), q(t)) - H(t, \bar{x}(t), \bar{u}(t), p(t) + r(t), q(t)) \\
& \frac{1}{2} (\Delta \sigma[t])'(P(t) + R(t))(\Delta \sigma[t]) \geq 0,
\end{aligned}
\]

for a.e. \(t \in [0, T]\), \(\mathbb{P}\)-a.s. and all \(v \in U\), where

\[
\begin{aligned}
\Delta \sigma[t] &= \sigma(t, \bar{x}(t), v) - \sigma(t, \bar{x}(t), \bar{u}(t)), \\
H(t, x, u, p, q) &= p \cdot b(t, x, u) + q \cdot \sigma(t, x, u) + \sum_{i=0}^M \Psi_i(t)f^i(t, x, u).
\end{aligned}
\]

The costates \((p, q)\) and \((P, Q)\) satisfy the standard adjoint equations associated with \(H(t, x, u, p, q)\) of first and second order with appropriate terminal conditions (see [49] or [35]).

The processes \(r(t)\) and \(R(t)\) are completely determined by the "bequest" terms \(h^i\) \((i = 0, \ldots, M)\) and the multipliers \((\lambda^0, \Psi^1, \ldots, \Psi^M)\).

To prove the principle, consider the metric on \(\mathcal{M}([0, T]; U)\):

\[
d_{\mathcal{M}}(u(\cdot), v(\cdot)) = m\{ (t, \omega) \in [0, T] \times \Omega : u(t, \omega) \neq v(t, \omega) \}
\]

where \(m\) is the product measure of Lebesgue measure and the probability \(\mathbb{P}\). It follows that \(d_{\mathcal{M}}\) is complete on \(\mathcal{M}([0, T]; U)\) ([49]), and so is the product metric of \(\mathcal{W} \equiv \mathbb{R}^n \times \mathcal{M}([0, T]; U)\). Therefore the stochastic optimal control problem is to minimize \(J^0(\omega)\) for \(\omega \equiv (x_0, u(\cdot))\) \(\in \mathcal{W}\) subject to \((11)\), which is just another example of Problem (1) in Section 1.

As mentioned above, a substantial part of the proof is to determine certain sequential strict derivatives \((z^0, z)\) of \((J^0, S)\) at \(w_0 = (\bar{x}_0, \bar{u}(\cdot))\) and then write inequality (3.2) in terms of conditions involving the Hamiltonian and the costate variables. This should lead to condition (B). Condition (A) follows from (3.1) and (3.3) of Theorem 1.

The difficulty of stochastic optimal controls is well-known. Specifically, if \((z^\varepsilon, u^\varepsilon(\cdot))\) is a variation of \((z, u(\cdot))\) with \(|z^\varepsilon - z| \approx \varepsilon\) and \(d_{\mathcal{M}}(u^\varepsilon(\cdot), u(\cdot)) \approx \varepsilon\). Let \(x^\varepsilon(\cdot)\) and \(x(\cdot)\) be the corresponding states. Then \(|x^\varepsilon(\cdot) - x(\cdot)| \approx \sqrt{\varepsilon}\) (not \(\approx \varepsilon\) as in the deterministic case). To overcome this difficulty, it is necessary to use costates of first order and second order to
represent sequential derivatives. The construction of the "spike" variations is similar to that of deterministic case; see [35], however, the estimates and variational formula in [35] and [22] are needed for estimates associated the costates of second order.

Of course, the set \( Q_P \) in the maximum principle can be replaced by a non-convex set once the general multiplier rules proposed in Section 3.2 are proved.

**Part II. Long-Term Plans (2008 and later)**

The long-term goals of this proposal include (1) to define subdifferentials for maps on metric spaces and use them to build a variational analysis for solving optimization problems; see Section 3.4 below, (2) to apply the results to more optimization problems; see Sections 3.5/3.6.

### 3.4. Subdifferentials and variational analysis on metric spaces

Directional derivatives and gradients/differentials of maps are dual concepts of variational analysis on Banach spaces. For maps on metric spaces, the notion of sequential derivatives has been defined in Section 2.1 as a generalization of directional derivatives. Now we need to define gradients/differentials for maps on metric spaces. Unfortunately, metric spaces usually do not have dual spaces to which the gradients/differentials of maps should belong. To mend this defect, it is natural (also sufficient for many applications) to consider metric spaces continuously embeddable into Banach spaces. Without loss of generality, we consider a subset \( W \) of a Banach space \( (X, \| \cdot \|_X) \), which is equipped with a metric \( d \) satisfying \( \| x - y \|_X \leq d(x, y) \) for all \( x, y \in W \). This inequality implies that the inclusion \( i : (W, d) \to (X, \| \cdot \|_X) \) is continuous. Now for \( F : W \to \mathbb{R} \cup \{ \infty \} \) and \( \varepsilon > 0 \), define the **\( \varepsilon \)-subdifferential** \( \partial_{\varepsilon} F(x; W) \) of \( F \) at \( x \in W \) as the following subset of \( X^* \):

\[
\partial_{\varepsilon} F(x; (W, d)) = \left\{ x^* \in X^* \mid \inf_{u \in (W, d) \to x} \frac{F(u) - F(x) - \langle x^*, u - x \rangle}{d(u, x)} \geq -\varepsilon \right\},
\]

where \( \langle \cdot, \cdot \rangle \) is the pairing between \( X^* \) and \( X \), and "\( u \in (W, d) \to x \)" means that "\( u \in W \) and \( d(u, x) \to 0 \)." The limit \( \partial F(\bar{x}; W) = \lim_{\varepsilon \to 0} \sup_{x \in (W, d) \to \bar{x}, \varepsilon \to 0} \partial_{\varepsilon} F(x; (W, d)) \) is called the **subdifferential** of \( F \) at \( \bar{x} \). These definitions generalize (and unify) existing notions of differentials of functions and normals of sets, as explained in following remarks.

**Remark 1.** If \( F \) is constant, then (14) reduces to \( \partial_{\varepsilon} F(x; (W, d)) = \left\{ x^* \in X^* \mid \lim_{u \in (W, d) \to x} \sup_{\varepsilon \to 0} \frac{\langle x^*, u-x \rangle}{d(u, x)} \leq \varepsilon \right\} \), which defines the set of **\( \varepsilon \)-normals** of the metric space \( (W, d) \) at \( x \), generalizing the notion of \( \varepsilon \)-normals of \( X \); see [32].

**Remark 2.** Assume that \( (X, \| \cdot \|_X) \) is an Asplund generated space (AGS), that is, there exists an Asplund subspace \( (W, \| \cdot \|_W) \) of \( (X, \| \cdot \|_X) \) such that \( W \) is dense in \( (X, \| \cdot \|_X) \)
and \(\|u\|_X \leq \|u\|_W\) for all \(u \in W\). Asplund generated spaces cover most of the spaces in applications yet they possess many nice properties of Asplund spaces. For every \(x \in X\), the set \(x + W\) is a metric space with metric \(d(u, v) = \|u - v\|_W\) for all \(u, v \in x + W\). For \(F : X \to \mathbb{R} \cup \{\infty\}\), the map \(x \in X \to \partial F(x; (x + W, d))\) defines a subdifferential of \(F\), which generalizes the subdifferential \(\partial_W F(x)\) defined in [12] but only for \(x \in W\).

**Remark 3.** Definition (14) easily extends to a set-valued map \(F\) from \((W, d)\) to another Banach space \((Y, \| \cdot \|_Y)\). Let \(\varepsilon > 0\) and \((x, y) \in \text{gph}(F) = \{x \in W : y \in F(x)\}\). The \(\varepsilon\)-subdifferential of \(F\) at \((x, y)\) is the set-valued map \(\tilde{\partial}_\varepsilon F((x, y); (W, d)) : Y^* \to X^*\) such that

\[
\tilde{\partial}_\varepsilon F((x, y); (W, d))(y^*) = \left\{ x^* \in X^* \mid \lim_{(u, v) \to (x, y)} \inf \frac{\langle y^*, v - y \rangle - \langle x^*, u - x \rangle}{d(u, x)} \geq -\varepsilon \right\},
\]

for each \(y^* \in Y^*,\) where \(\langle x^*, u - x \rangle\) and \(\langle y^*, v - y \rangle\) are pairings between \(X^*\) and \(X, Y^*\) and \(Y\), respectively. If \(F\) is single-valued, then definition (15) simplifies to

\[
\tilde{\partial}_\varepsilon F(x; W)(y^*) = \left\{ x^* \in X^* \mid \lim_{u \to x} \inf \frac{\langle y^*, F(u) - F(x) \rangle - \langle x^*, u - x \rangle}{d(u, x)} \geq -\varepsilon \right\}.
\]

**Remark 4.** By adding \(\|F(u) - F(x)\|_W, \|v - y\|_W\) and \(\|F(u) - F(x)\|_W\) to \(d(u, x)\) in (14)-(16), we obtain the definitions of the corresponding geometric \(\varepsilon\)-subdifferentials of \(F\).

**Remark 5.** Various notions of tangent cones can be defined for \(W\) as a subset of \((X, \| \cdot \|_X)\).

With these notions, a variational analysis for solving optimization problems like Problem (1) in Section 1 can be built with the following two plans.

(1) First consider the case \(W \subset (X, \| \cdot \|_X)\) with \(d(x, y) = \|x - y\|_X\) and focus on the relationships among subdifferentials and sequential derivatives of maps on \(W\) and the geometry (tangents/normals) of \(W\). These relationships lead to direct applications of the multiplier rules in Section 3.2 to the optimization problem (4) in Section 2 with geometric constraints.

(2) Then consider an Asplund generated space \((X, \| \cdot \|_X)\) and a subset \(W\) equipped with metric \(d\) satisfying \(\|x - y\|_X \leq d(x, y)\) for \(x, y \in W\). This case is especially important because many spaces such as \(C[a, b]\) in applications are Asplund generated spaces but not Asplund spaces. The goal is to extend the elegant variational analysis of Asplund spaces ([32, Chapters 2/3] [12] [13]) to Asplund generated spaces. Consequently, the multiplier rules in Section 3.2 for Problem (1) in Section 1 can be further generalized to the most general cases.

In summary, by embedding metric spaces into Banach spaces, various notions of tangents and normals for sets and gradients/subdifferentials for maps can be defined properly, and a variational analysis for solving optimization problems on metric spaces can be established. Although a full-scale variational analysis on metric spaces is beyond the scope of this proposal,
the framework of such analysis has emerged from the initial works [27], [29] and [34] of this proposal. Stimulated by this research project, many researches on variational analysis on metric spaces will appear in years to come, generalizing the rich theories of variational analysis on Banach spaces, including [4], [5], [6], [8], [9], [24], [32], [33], and [42].

3.5. Multi-objective optimizations on metric spaces

My another long-term plan will be to establish multiplier rules for *multi-objective optimization* problems on metric spaces. Such multiplier rules can be applied to optimal controls with multiple objectives to derive "equilibrium principles" for solutions. For applications, I will focus on *stochastic* optimal controls with multiple objectives, especially *stochastic games*. The paper [35] has obtained the basic estimates for vector-valued objective functionals (associated with general stochastic differential equations), which are needed for *multi-objective* stochastic optimal controls. This project is closely related to the investigator's another research interest on differential games in [36] [37].

3.6. Further applications

As further applications, I will look into the following four classes of optimal controls.

(i) stochastic controls with free end points

(ii) stochastic optimal control with controlled partial differential equations

(iii) stochastic optimal controls with delayed equations

(iv) optimal controls with backward stochastic equations

As shown in [27], multiplier rules on metric spaces may significantly simplify or unify the existing proofs of maximum principles or strengthen existing conclusions. The literature of existing works on these optimal controls are extensive, which will not be listed here. The main goal of this endeavor is *not to reprove* existing results, but to derive necessary conditions for solutions in more general or new settings, especially in the setting with stochastic controls, which are more and more important in applications, but much fewer results are known for them.

Part III: How the Plans Will Be Carried Out.

The investigator will be responsible for carrying out all of the proposed research. He will continue to work with his current collaborators Professor Michael McAsey at Bradley University (on part of Sections 3.1 and 3.6), Professor Jiongmin Yong at University of Central Florida (on part of Sections 3.3 and 3.6). Consultation is being sought with Professor Boris Mordukhovich at Wayne State University (on Sections 3.2, 3.4 and 3.5).
Section 4. Broader Impacts of the Proposal

4.1. On the discipline of mathematics

Developing a variational analysis on metric spaces would be a dream of variational analysts. On one hand, there are beautiful and very useful variational analyses (including smooth/nonsmooth and linear/nonlinear analysis) on Euclidean, Hilbert and Banach spaces. On the other hand, many real-life mathematical problems naturally exist on metric spaces, and metric spaces do not have the structures (of a Banach space) for defining differentiation (either in classical or generalized senses). This research develops a variational analysis on metric spaces by introducing the notions of sequential derivatives (defined in Section 2). While sequential derivatives do not have all of the nice properties of classical derivatives (as expected), they are right notions for building a variational analysis on metric spaces (described in Section 3.1), which aims at solving optimization/equilibrium problems. The applications to optimal controls (deterministic case in [27] and stochastic case in Section 3.3 of this proposal) show the strength and convenience of the variational analysis to be developed in this research.

This analysis complements other analyses ([1], [2], [21], [43], [44], [45]) developed on metric spaces, whose focus is on geometric aspects of metric spaces. The general multiplier rules for Problem (1) (in Section 1) involve the variational analysis on the metric space $\mathcal{W}$ and nonsmooth analysis on the Banach space $Z$ (see Sections 3.2 and 3.4). Therefore, the proposed research will stimulate further studies of analyses on both metric spaces and Banach spaces.

4.2. On other disciplines

At Bradley University the investigator has opportunities to work with some economists and engineers on optimal control problems arising from economy, finance and engineering. These experiences partially motivated him to search for general multiplier rules on metric spaces that are applicable to various optimization problems. The variational analysis and multiplier rules developed in this research can be used by applied mathematicians and scientists to derive necessary conditions for solutions of their problems. Since the existence of multipliers is guaranteed, they can focus on determining appropriate sequential derivatives (which has to be done case by case) in order to derive their results.

4.3. On other aspects

The proposed research also has significant impacts on research and educational environments at the investigator’s home institution. Please see RUI Impact Statement included in the Supplementary Documents.
References Cited
in
"Variational Analysis for Optimizations on Metric Spaces and Applications"

(The investigator's papers are available at http://hilltop.bradley.edu/~mou/moupaper.html)


