Pontryagin’s Maximum Principle for a Advection-Diffusion-Reaction Equation

Youjun Xu$^{1,2}$, Cuie Xiao$^3$, Hui Zhu$^1$

$^1$School of Mathematics and Physics, University of South China, Hengyang, China
$^2$School of Mathematical Sciences, Fudan University, Shanghai, China
$^3$Department of Mathematics and Computation Sciences, Hunan City University, Yiyang, China

Email: youjunxu@163.com, xiaocuie@163.com

Received July 2, 2012; revised November 19, 2012; accepted November 26, 2012

ABSTRACT

In this paper we investigate optimal control problems governed by an advection-diffusion-reaction equation. We present a method for deriving conditions in the form of Pontryagin’s principle. The main tools used are the Ekeland’s variational principle combined with penalization and spike variation techniques.

Keywords: Optimal Control; Pontryagin’s Maximum Principle; State Constraint

1. Introduction

Consider the following controlled advection convection diffusion equations:

\[
\begin{cases}
-\nabla \cdot (\mu \nabla y) + \beta \nabla y + \sigma y = f(x,u) \text{ in } \Omega, \\
y = 0 \text{ on } \partial \Omega,
\end{cases}
\]

where $\Omega \subset \mathbb{R}^n$, $n \geq 2$ is a convex bounded domain with a smooth boundary $\partial \Omega$, the diffusity $\mu \in L^\infty(\Omega)$ with $\mu \geq \mu_0 > 0$ a.e. in $\Omega$, the reaction $\sigma \in L^\infty(\Omega)$ with $\sigma \geq \sigma_0 > 0$, and the advective field $\beta \in \left( L^\infty(\Omega) \right)^2$, with $\nabla \beta \in L^\infty(\Omega)$ and $-\frac{1}{2} \nabla \beta + \sigma \geq 0$ a.e. in $\Omega$ are assigned functions. Here $f: \Omega \times U \to \mathbb{R}$, with $U$ being a separable metric space. Function $u(\cdot)$, called a control, is taken from the set

\[U = \{ w: \Omega \to U | w(\cdot) \text{ is measurable} \} .\]

Under some mild conditions, for any $u(\cdot) \in U$, (1.1) admits a unique weak solution $y(\cdot) = y(\cdot; u(\cdot))$, which is called the state(corresponding to the control $u(\cdot)$). The performance of the control is measured by the cost functional

\[J(u(\cdot)) = \int_\Omega f^0(x,y(x),u(x))dx.\]

for some given map $f^0: \Omega \times U \to \mathbb{R}$. Our optimal control problem can be stated as follows.

Problem (C). Find a $u(\cdot) \in U$ such that

\[J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in U} J(u(\cdot)).\]  

And the state constraint of form:

\[F(y) \in Q.\]

In this paper, we make the following assumptions.

(H1) Set $\Omega \subset \mathbb{R}^n$, $n \geq 2$ is a convex bounded domain with a smooth boundary $\partial \Omega$.

(H2) Set $U$ is a separable metric space.

(H3) The function $f: \Omega \times U \to \mathbb{R}$ has the following properties: $f(\cdot; u)$ is measurable on $\Omega$, and $f(x,\cdot)$ continuous on $\Omega \times U$ and for any $R > 0$, a constant $M_R > 0$, such that $|f(x,u)| \leq M_R, \forall (x,u) \in \Omega \times U$.

(H4) Function $f^0(x,y,v)$ is measurable in $x$ and continuous in $(y,v) \in R \times U$ for almost all $x \in \Omega$. Moreover, for any $R > 0$, there exists a $K_R > 0$ such that

\[|f^0(x,y,v)| + |f^0(x,y,v)| \leq K_R, \quad \text{a.e. } x,v \in \Omega \times U, |y| \leq R.\]

(H5) $\Omega$ is a Banach space with strictly convex dual $\Omega^*$, $F: W_0^{1,p}(\Omega) \to X$ is continuously Fréchet differentiable, and $Q \subset X$ is closed and convex set.

(H6) $F(\bar{y})D_y - Q$ has finite codimensionality in $X$ for some $r > 0$, where $D_y = \{ z \in X : \|z\|_X \leq r \}$.

Definition 1.1 (see [1]) Let $X$ is a Banach space and $X_0$ is a subspace of $X$. We say that $X_0$ is finite codimensional in $X$ if there exists $x_1, x_2, \cdots, x_n \in X$ such that

\[\|x - \sum_{i=1}^n a_i x_i\|_X = \inf_{a \in \mathbb{R}^n} \|x - \sum_{i=1}^n a_i x_i\|_X = 0.\]
span \{X_0, x_1, \cdots, x_n\} = \text{the space spanned by}
\{X_0, x_1, \cdots, x_n\} = X.

A subset \( S \) of \( X \) is said to be finite codimensional in \( X \) if for some \( x_0 \in S \), \( \text{span} \{S - \{x_0\}\} \) is a finite codimensional subspace of \( X \) and \( coS \) the closed convex hull of \( S - \{x_0\} \) has a nonempty interior in this subspace.

**Lemma 1.2.** Let \((H1) - (H3)\) hold. Then, for any \( u(\cdot) \in U \), (1.1) admits a unique weak solution
\[
y(\cdot) \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega). \tag{1.6}
\]
Furthermore, there exists a constant \( K > 0 \), independent of
\[
u(\cdot) \in U, \|y(\cdot)\|_{W_0^{1,p}(\Omega) \cap L^\infty(\Omega)} \leq K \tag{1.6}
\]
The weak solution \( y \in V = H_0^1(\Omega) \) of the state Equation (1.1) is determined by
\[
a(y, v) = (f, v), \forall v \in V.
\]
using the bilinear form \( a: V \times V \rightarrow R \) given by
\[
a(y, v) = \int_{\Omega} \mu \nabla y \nabla v \, dx + \int_{\Omega} \beta \nabla y \nabla v \, dx + \int_{\Omega} \sigma y v, \forall v \in V.
\]
Existence and uniqueness of the solution to (1.1) follow from the above hypotheses on the problem data (see [2]). Let \( A_{ad} \) be the set of all pairs \((y(\cdot), u(\cdot))\) satisfying (1.1) and (1.4) is called an admissible set. Any \((y, u) \in A_{ad}\) is called an admissible pair. The pair \((\overline{y}(\cdot), \overline{u}(\cdot)) \in A_{ad}\) mover satisfies \( J(\overline{u}(\cdot)) \leq J(u) \) for all \((y, u) \in A_{ad}\) is called an optimal pair. If it exists, refer to \( \overline{y} \) and \( \overline{u} \) as an optimal state and control, respectively.

Now, let \((\overline{y}, \overline{u})\) be an optimal pair of Problem (C). Let \( z = z(\cdot; u(\cdot)) \in W_0^{1,p}(\Omega) \) be the unique solution of the following problem:
\[
\begin{aligned}
-\nabla \cdot (\mu \nabla z) + \nu \nabla z + \sigma z &= f(x, u) - f(x, \overline{u}) \text{ in } \Omega, \\
z &= 0 \text{ on } \partial \Omega. 
\end{aligned} \tag{1.7}
\]
And define the reachable set of variational system (1.7)
\[
R = \{z(\cdot; u(\cdot)) | u(\cdot) \in U\}. \tag{1.8}
\]
Now, let us state the first order necessary conditions of an optimal control to Problem (C) as follows.

**Theorem 1.3.** (Pontryagin’s maximum principle) Let \((H1) - (H6)\) hold. Let \((\overline{y}(\cdot), \overline{u}(\cdot))\) be an optimal pair of Problem (C). Then there exists a triplet
\[
(\Psi^0, \Psi, \phi) \in R \times W_0^{1,p} \times X^* \text{ with } (\Psi^0, \phi) \neq 0,
\]
such that
\[
\langle \phi, \eta - F(\overline{y}) \rangle_{X^*, X} \leq 0, \forall \eta \in Q. \tag{19}
\]
\[
\begin{align*}
\bar{d}\left(u, u^e\right) & \leq \sqrt{\varepsilon}, \\
J_{\varepsilon}(\bar{u}) - J_{\varepsilon}\left(u^e\right) & \geq -\varepsilon \bar{d}\left(\bar{u}, u^e\right), \forall \bar{u} \in U.
\end{align*}
\]

Let \(v \in U\) and \(\varepsilon > 0\) be fixed and let \(y^e = y(\cdot; u^e)\), we know that for any \(\rho \in (0, 1)\), there exists a measurable set \(E^{\varepsilon}_\rho \subset \Omega\) with the property \(|E^{\varepsilon}_\rho| = \rho|\Omega|\), such that if we define

\[
u^{\varepsilon}_\rho(x) = \begin{cases} u^e(x), & \text{if } x \in \Omega \setminus E^{\varepsilon}_\rho, \\ v(x), & \text{if } x \in E^{\varepsilon}_\rho. \end{cases}
\]

and let \(y^{\varepsilon}_\rho = y(\cdot; u^{\varepsilon}_\rho)\) be the corresponding state, then

\[
\begin{align*}
y^{\varepsilon}_\rho &= y^e + \rho z^{\varepsilon} + r^{\varepsilon}, \\
J_{\varepsilon}\left(u^{\varepsilon}_\rho(x)\right) &= J(u^e) + \rho z^{\varepsilon} + r^{\varepsilon},
\end{align*}
\]

where \(z^{\varepsilon}\) and \(z^{\varepsilon}_\rho\) satisfying the following

\[
\begin{align*}
-\nabla \left(\mu(\nabla z^{\varepsilon})\right) + \beta \nabla z^{\varepsilon} + \sigma z^{\varepsilon} = f(x,v) - f(x,u) \quad &\text{in } \Omega, \\
z^{\varepsilon} = 0 \quad &\text{on } \partial \Omega, \\
z^{\varepsilon}_\rho = \frac{1}{\Omega} \int_\Omega \left[J_{\varepsilon}(x, y, u^e) + h^{\varepsilon}(x)\right] dx
\end{align*}
\]

with

\[
\begin{align*}
h^{\varepsilon}(x) &= f^0(x, y^e, v) - f^0(x, y^e, u^e), \\
\lim_{\rho \to 0} \frac{1}{\rho} \left\|F^{\varepsilon}_\rho\right\|_{W^{1,\rho}_0(\Omega)} &= \lim_{\rho \to 0} \frac{1}{\rho} \left\|F^{\varepsilon}_\rho\right\|_{W^{1,\rho}_0(\Omega)} = 0.
\end{align*}
\]

We take \(\bar{u} = u^{\varepsilon}_\rho\). It follows that

\[
\begin{align*}
&\sqrt{\varepsilon}|\Omega| \leq J_{\varepsilon}(u^{\varepsilon}_\rho) - J(u^e) \\
= & \frac{1}{\rho} \left(J_{\varepsilon}(u^{\varepsilon}_\rho) + J_{\varepsilon}\left(u^e\right)\right) \\
+ & \frac{1}{\rho} \left[J_{\varepsilon}(u^{\varepsilon}_\rho) + \varepsilon\right] - \frac{1}{\rho} \left[J(u^e) + \varepsilon\right] \\
+ & \frac{d_{\varepsilon}(F(y^{\varepsilon}_{\rho})) - d_{\varepsilon}(F(y^{\varepsilon}))}{\rho} \\
\to & \frac{\left(J(u^e) + \varepsilon\right)}{J_{\varepsilon}\left(u^e\right)} z^{\varepsilon}_\rho \\
& + \frac{d_{\varepsilon}(F(y^{\varepsilon}_{\rho}))}{J_{\varepsilon}\left(u^e\right)} \xi\left(F'(y^{\varepsilon})z^{\varepsilon}\right) \quad \text{as } \rho \to 0.
\end{align*}
\]

where

\[
\xi\left(F'(y^{\varepsilon})z^{\varepsilon}\right) = \begin{cases} \nabla d_{\varepsilon}\left(F(y^{\varepsilon})\right) & \text{if } F(y^{\varepsilon}) \notin Q, \\
0 & \text{if } F(y^{\varepsilon}) \in Q.
\end{cases}
\]

Next, we define \((\phi^{\varepsilon,\rho}, \varphi) \in [0,1] \times X^*\) as follows:

\[
\phi^{\varepsilon,\rho} = \frac{J(u^e) + \varepsilon}{J_{\varepsilon}(u^e)} \xi_{\varepsilon}\left(F'(y^{\varepsilon})z^{\varepsilon}\right)
\]

By (2.1) and chapter 4 of [8], (2.8) becomes

\[
\phi^{\varepsilon,\rho} \left| + \phi^{\varepsilon,\rho} \right|_{z^{\varepsilon}} = 1.
\]

On the other hand, by the definition of the subdifferential, we have

\[
\left\langle \phi^{\varepsilon,\rho}, \eta - F(y^{\varepsilon}) \right\rangle \leq 0, \forall \eta \in Q
\]

Next, from the first relation in (2.3) and by some calculations, we have

\[
\|y^{\varepsilon} - \bar{y}\|_{W^{1,\rho}(\Omega)} \to 0, \quad \text{as } \varepsilon \to 0.
\]

Consequently,

\[
\lim_{\varepsilon \to 0} \left\|\phi^{\varepsilon,\rho} - \phi^e\right\|_{V^*} = 0.
\]

From (2.5) and (2.6), we have

\[
\begin{align*}
z^{\varepsilon}_\rho & \to z \quad \text{in } W^{1,\rho}_0(\Omega), \\
z^{\varepsilon}_\rho & \to z^0 \quad \text{as } \rho \to 0,
\end{align*}
\]

where \(z\) is the solution of system (1.7) and

\[
\begin{align*}
z^0 &= \int_\Omega f^0(x, y^e, v) z(x) dx \\
&+ \int_\Omega \left[f^0(x, y^e, v) - f^0(x, y^e, u^e)\right] dx.
\end{align*}
\]

From (2.10), (2.12) and (2.15), we have

\[
\phi^{\varepsilon,\rho} z(v) + \left\langle \phi^e, F'(y^e)z(v) - \eta + F'(\bar{y})\right\rangle \\
\geq -\delta_{\varepsilon}, \quad \forall v \in U, \eta \in Q.
\]

Now, let
\[ \psi^0 = -\varphi^0 \in [-1, 0]. \]

Then
\[ (\psi^0, \varphi) \neq 0. \]

Then we have
\[ \varphi^0 z^0(v) + \langle \varphi, \eta - F'(\psi) \rangle - F'(\psi)^* \varphi, z(v) \rangle \geq 0, \quad \forall \eta \in U, \quad \forall \varphi \in Q. \] 

(2.19)

Take \( \nu = \bar{\nu} \), we obtain (1.9).

Next, we let \( \eta = F(\psi) \) to get
\[ \psi^0 \varphi(v) - \langle F'(\psi)^* \varphi, z(v) \rangle \leq 0 \quad \forall \nu \in U. \] 

(2.20)

Because \( F'(\psi)^* \varphi \in W^{-1,r}(\Omega) \), for the given \( \psi^0 \), there exists a unique solution \( \psi \in W^{1,r}(\Omega) \) of the adjoint Equation (1.10). Then, from (1.6), (2.16), and (2.2), we have
\[ 0 \geq \psi^0 z^0(v) - \langle F'(\psi)^* \varphi, z(v) \rangle \]
\[ = \int_\Omega f^0(x, \bar{\psi}, \bar{\nu}) z(x) \, dx \]
\[ + \int_\Omega \left[ f^0(x, \bar{\psi}, v) - f^0(x, \bar{\psi}, \bar{\nu}) \right] \, dx \]
\[ + \left\{ - \nabla \left( \mu \nabla \bar{\psi} \right) - \nabla \left( \beta \bar{\psi} \right) + \sigma \bar{\psi} - \psi^0 f^0(x, \bar{\psi}, \bar{\nu}) , z \right\} \]
\[ = \int_\Omega \left\{ \psi^0 \left[ f^0(x, \bar{\psi}, v) - f^0(x, \bar{\psi}, \bar{\nu}) \right] \right\} \, dx \]
\[ + \left\{ \psi^0 f(x, \nu) - f(x, \bar{\nu}) \right\} \, dx \]
\[ = \int_\Omega \left\{ H(x, \bar{\psi}(x), v(x), \psi^0, \psi(x)) \right\} \, dx \]
\[ - \left\{ H(x, \bar{\psi}(x), \bar{\nu}(x), \psi^0, \psi(x)) \right\} \, dx \]

There, (1.11) follows. Finally, by (1.10), if \( (\psi^0, \psi) = 0 \), then \( F'(\psi)^* \varphi = 0 \). Thus, in the case where
\[ N \left( F'(\psi)^* \right) = \{ 0 \}, \]

we must have \( (\psi^0, \psi) \neq 0 \), because \( (\psi^0, \varphi) \neq 0 \).

### 3. Conclusion

We have already attained Pontryagin’s Maximum Principle for the advection-diffusion-reaction equation. It seems to us that this method can be used in treating many other relevant problems.

### REFERENCES


