Applied Mathematics, 2012, 3, 1888-1891
http://dx.doi.org/10.4236/am.2012.312258 Published Online December 2012 (http://www.SciRP.org/journal/am)

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 a 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:

$$\begin{cases} -\nabla\cdot(\mu\nabla y)+\beta\nabla y+\sigma y = f(x,u) \text{ in } \Omega, \\ y=0 \text{ on } \partial\Omega. \end{cases} \tag{1.1}$$

where $\Omega\subset R^n$, $n\geq 2$ 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 Ω, 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 Ω are assigned functions. Here $f:\Omega\times U\to R$, with U being a separable metric space. Function $u(\cdot)$, called a control, is taken from the set

$$\mathcal{U}=\left\{w:\Omega\to U \,\middle|\, w(\cdot) \text{ is measurable}\right\}.$$

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

$$J(u(\cdot))=\int_\Omega f^0(x,y(x),u(x))\,\mathrm{d}x. \tag{1.2}$$

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

Problem (C). Find a $\bar{u}(\cdot)\in\mathcal{U}$ such that

$$J(\bar{u}(\cdot))=\inf_{u(\cdot)\in\mathcal{U}} J(\bar{u}(\cdot)). \tag{1.3}$$

And the state constraint of form:

$$F(y)\in Q. \tag{1.4}$$

In this paper, we make the following assumptions.

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

(H2) Set U is a separable metric space.

(H3) The function $f:\Omega\times U\to R$ has the following properties: $f(\cdot;u)$ is measurable on Ω, 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

$$\begin{aligned} &\left|f^0(x,y,v)\right|+\left|f_y^0(x,y,v)\right|\leq K_R, \\ &a.e.\ x,v\in\Omega\times U, |y|\leq R. \end{aligned} \tag{1.5}$$

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

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

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

*This work was partially supported by National Natural Science Foundation of China (Grant No. 11126170) and Hunan Provincial Natural Science Foundation of China (Grant No. 11JJ4006) and Ph.D startup Foundation of University of South China (Grant No 5-2011-XQD-008).

$\text{span}\{X_0, x_1, \cdots, x_n\}$ = 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 $x_0 \in S$, $\text{span}(S - \{x_0\})$ the closed subspace spanned by $\{x - x_0 | x \in S\}$ is a finite codimensional subspace of X and $\overline{co}S$ 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

$$y(\cdot) \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega).$$

Furthermore, there exists a constant $K > 0$, independent of

$$u(\cdot) \in U, \|y(\cdot)\|_{W_0^{1,p}(\Omega) \cap L^\infty(\Omega)} \le K \tag{1.6}$$

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 $a : V \times V \to R$ given by

$$a(y, v) = \int_\Omega \mu \nabla y \nabla v \mathrm{d}x + \int_\Omega \beta \nabla y v \mathrm{d}x + \int_\Omega \sigma y v \mathrm{d}x, \ \forall\, v \in V.$$

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 $(\bar{y}(\cdot), \bar{u}(\cdot)) \in A_{ad}$, moveover satisfies $J(\bar{u}(\cdot)) \le J(u)$ for all $(y, u) \in A_{ad}$ is called an optimal pair. If it exists, refer to $\bar{y}$ and $\bar{u}$ as an optimal state and control, respectively.

Now, let $(\bar{y}, \bar{u})$ 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:

$$\begin{cases} -\nabla \cdot (\mu \nabla z) + \beta \nabla z + \sigma z = f(x, u) - f(x, \bar{u}) \text{ in } \Omega, \\ z = 0 \text{ on } \partial\Omega. \end{cases} \tag{1.7}$$

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 $(\bar{y}(\cdot), \bar{u}(\cdot))$ be an optimal pair of Problem (C). Then there exists a triplet

$$(\Psi^0, \Psi, \phi) \in R \times W_0^{1,p'} \times X^* \text{with} (\psi^0, \varphi) \neq 0,$$

such that

$$\langle \varphi, \eta - F(\bar{y}) \rangle_{X^*, X} \le 0, \forall \eta \in Q. \tag{19}$$

$$\begin{cases} -\nabla \cdot (\mu \nabla \psi) - \nabla(\beta \psi) + \sigma \psi \\ = \psi^0 f_y^0(x, y, u) - F(y)^* \varphi \text{ in } \Omega, \\ \psi = 0 \text{ on } \partial\Omega. \end{cases} \tag{1.10}$$

$$\begin{aligned} & H(x, \bar{y}(x), \bar{u}(x), \psi^0, \psi(x)) \\ & = \max_{u(\cdot) \in U} H(x, \bar{y}(x), u(x), \psi^0, \psi(x)) \ a.e.\ x \in \Omega. \end{aligned} \tag{1.11}$$

where

$$\begin{aligned} & H(x, \bar{y}(x), u(x), \psi^0, \psi(x)) \\ & = \psi^0 f^0(x, \bar{y}, u) + \psi(x) f(x, u), \end{aligned}$$

(1.9), (1.10), and (1.11) are called the transversality condition, the adjoint system(along the given optimal pair), and the maximum condition, respectively.

Many authors (Dede [3], Yan [4], Becker [5], Stefano [6], Collis [7]) have already considered control problems for convection-diffusion equations from theoretical or numerical point of view. In the work mentioned above, the control set is convex. However, in many practical cases, the control set can not convex. This stimulates us to study Problem (C). To get Pontryagin's Principle, we use a method based on penalization of state constraints, and Ekeland's principle combined with diffuse perturbations [8].

In the next section, we will prove Pontryagin's maximum principle of optimal control of Problem (C).

2. Proof of the Maximum Principle

This section is devoted to the proof of the maximum principle.

Proof of Theorem 1.3. Firstly, let

$$\bar{d}(u(\cdot), \bar{u}(\cdot)) = \left|\{x \in \Omega | u(\cdot) \neq \bar{u}(\cdot)\}\right|,$$

where $|D|$ is the Lebesgue measure of $D \subseteq \Omega$. We can easily prove that (u, ρ) is a complete metric space. Let $(\bar{y}, \bar{u})$ be anoptimal pair of Problem (C). For any $y(\cdot; u)$ be the corresponding state, emphasizing the dependence on the control. Without loss of generality, we may assume that $J(\bar{u}) = 0$. For any $\varepsilon > 0$, define

$$J_\varepsilon(u) = \left\{\left[\left(J(u) + \varepsilon\right)^+\right]^2 + d_Q^2\left(F\left(y(\cdot, u)\right)\right)\right\}^{1/2} \tag{2.1}$$

where $d_Q = \inf_{\bar{x} \in Q} |x - \bar{x}|$, and $\bar{u}$ is an optimal control. Clearly, this function is continuous on the (complete) metric space $(U, \bar{d})$. Also, we have

$$\begin{cases} J_\varepsilon(u) > 0, \forall u \in U, \\ J_\varepsilon(\bar{u}) = \varepsilon \le \inf\limits_U J(u) + \varepsilon. \end{cases} \tag{2.2}$$

Hence, by Ekeland's variational principle, we can find a $u^\varepsilon \in U$, such that

$$\begin{cases} \bar{d}\left(u,u^{\varepsilon}\right) \le \sqrt{\varepsilon}, \\ J_{\varepsilon}\left(\hat{u}\right) - J_{\varepsilon}\left(u^{\varepsilon}\right) \ge -\sqrt{\varepsilon}\bar{d}\left(\hat{u},u^{\varepsilon}\right), \forall \hat{u} \in U. \end{cases} \tag{2.3}$$

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

$$u_{\rho}^{\varepsilon}(x) = \begin{cases} u^{\varepsilon}(x), \text{ if } x \in \Omega / E_{\rho}^{\varepsilon}, \\ v(x), \text{ if } x \in E_{\rho}^{\varepsilon}. \end{cases}$$

and let $y_{\rho}^{\varepsilon} = y\left(\cdot;u_{\rho}^{\varepsilon}\right)$ be the corresponding state, then

$$\begin{cases} y_{\rho}^{\varepsilon} = y^{\varepsilon} + \rho z^{\varepsilon} + r_{\rho}^{\varepsilon}, \\ J_{\varepsilon}\left(u_{\rho}^{\varepsilon}(x)\right) = J\left(u^{\varepsilon}\right) + \rho z^{0,\varepsilon} + r_{\rho}^{0,\varepsilon}, \end{cases} \tag{2.4}$$

where z^{ε} and $z^{0,\varepsilon}$ satisfying the following

$$\begin{cases} -\nabla \cdot \left(\mu\left(\nabla z^{\varepsilon}\right)\right) + \beta \nabla z^{\varepsilon} + \sigma z^{\varepsilon} \\ = f(x,v) - f(x,u) \text{ in } \Omega, \\ z^{\varepsilon} = 0 \text{ on } \partial\Omega. \end{cases} \tag{2.5}$$

$$z^{0,\varepsilon} = \int_{\Omega}\left[f_y^0\left(x,y,u^{\varepsilon}\right) z + h^{0,\varepsilon}(x)\right]\mathrm{d}x \tag{2.6}$$

with

$$\begin{cases} h^{0,\varepsilon}(x) = f^0\left(x,y^{\varepsilon},v\right) - f^0\left(x,y^{\varepsilon},u^{\varepsilon}\right) \\ \lim\limits_{\rho\to 0}\frac{1}{\rho}\left\|r_{\rho}^{\varepsilon}\right\|_{W_0^{1,p}} = \lim\limits_{\rho\to 0}\frac{1}{\rho}\left|r_{\rho}^{0,\varepsilon}\right| = 0. \end{cases} \tag{2.7}$$

We take $\hat{u} = u_{\rho}^{\varepsilon}$. It follows that

$$\begin{aligned} -\sqrt{\varepsilon}|\Omega| &\le \frac{J_{\varepsilon}\left(u_{\rho}^{\varepsilon}\right) - J\left(u^{\varepsilon}\right)}{\rho} \\ &= \frac{1}{J_{\varepsilon}\left(u_{\rho}^{\varepsilon}\right) + J_{\varepsilon}\left(u^{\varepsilon}\right)} \\ &\cdot \left\{ \frac{\left[\left(J\left(u_{\rho}^{\varepsilon}\right)+\varepsilon\right)^{+}\right]^2 - \left[\left(J\left(u^{\varepsilon}\right)+\varepsilon\right)^{+}\right]^2}{\rho} \right. \\ &\left. + \frac{d_Q^2\left(F\left(y_{\rho}^{\varepsilon}\right)\right) - d_Q^2\left(F\left(y^{\varepsilon}\right)\right)}{\rho} \right\} \\ &\to \frac{\left(J\left(u^{\varepsilon}\right)+\varepsilon\right)^{+}}{J_{\varepsilon}\left(u^{\varepsilon}\right)} z^{0,\varepsilon} \\ &+ \left\langle \frac{d_Q\left(F\left(y^{\varepsilon}\right)\right)\xi_{\varepsilon}}{J_{\varepsilon}\left(u^{\varepsilon}\right)}, F'\left(y^{\varepsilon}\right) z^{\varepsilon} \right\rangle \text{ as } \rho \to 0. \end{aligned} \tag{2.8}$$

where

$$\xi_{\varepsilon} = \begin{cases} \nabla d_Q\left(F\left(y^{\varepsilon}\right)\right) & \text{if } F\left(y^{\varepsilon}\right) \notin Q, \\ 0 & \text{if } F\left(y^{\varepsilon}\right) \in Q. \end{cases}$$

$\nabla d_Q(\cdot)$ denotes the subdifferential of $d_Q(\cdot)$.

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

$$\varphi^{0,\varepsilon} = \frac{\left(J\left(u^{\varepsilon}\right)+\varepsilon\right)^{+}}{J_{\varepsilon}\left(u^{\varepsilon}\right)}, \varphi^{\varepsilon} = \frac{d_Q\left(F\left(y^{\varepsilon}\right)\right)\xi_{\varepsilon}}{J_{\varepsilon}\left(u^{\varepsilon}\right)} \tag{2.9}$$

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

$$-\sqrt{\varepsilon}|\Omega| \le \varphi^{0,\varepsilon} z^{0,\varepsilon} + \left\langle \varphi^{\varepsilon}, F'\left(y^{\varepsilon}\right) z^{\varepsilon} \right\rangle, \tag{2.10}$$

$$\left|\varphi^{0,\varepsilon}\right|^2 + \left\|\varphi^{\varepsilon}\right\|_{X^*}^2 = 1. \tag{2.11}$$

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

$$\left\langle \phi^{\varepsilon}, \eta - F\left(y^{\varepsilon}\right) \right\rangle \le 0, \forall \eta \in Q \tag{2.12}$$

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

$$\left\|y^{\varepsilon} - \bar{y}\right\|_{W_0^{1,p}(\Omega)} \to 0, \text{ as } \varepsilon \to 0. \tag{2.13}$$

Consequently,

$$\lim_{\varepsilon\to 0}\left\|F'\left(y^{\varepsilon}\right) - F'(\bar{y})\right\|_{\ell\left(W_0^{1,p}(\Omega)\right)} = 0 \tag{2.14}$$

From (2.5) and (2.6), we have

$$\begin{cases} z^{\varepsilon} \to z \text{ in } W_0^{1,p}(\Omega), \\ z^{0,\varepsilon} \to z^0 \end{cases} \text{as } \varepsilon \to 0. \tag{2.15}$$

where z is the solution of system (1.7) and

$$\begin{aligned} z^0 &= \int_{\Omega} f_y^0\left(x,\bar{y},\bar{u}\right) z(x)\mathrm{d}x \\ &+ \int_{\Omega}\left[f^0\left(x,\bar{y},v\right) - f^0\left(x,\bar{y},\bar{u}\right)\right]\mathrm{d}x \end{aligned} \tag{2.16}$$

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

$$\begin{aligned} &\varphi^{0,\varepsilon} z(v) + \left\langle \varphi, F'(\bar{y}) z(\cdot;v) - \eta + F(\bar{y}) \right\rangle \\ &\ge -\delta_{\varepsilon}, \forall v \in U, \eta \in Q. \end{aligned} \tag{2.17}$$

with $\delta_{\varepsilon} \to 0$ as $\varepsilon \to 0$. Because $F'(\bar{y})D_r - W$ has finite condimensionality in X, we can extract some subsequence, still denoted by itself, such that

$$\left(\varphi^{0,\varepsilon},\varphi^{\varepsilon}\right) \xrightarrow{\ *\ } \left(\varphi^0,\varphi\right) \ne 0$$

From (2.17), we have

$$\begin{aligned} &\varphi^0 z^0(v) + \left\langle \varphi, F'(\bar{y}) z(\cdot;v) - \eta + F(\bar{y}) \right\rangle \ge 0, \\ &\forall v \in U, \eta \in Q \end{aligned} \tag{2.18}$$

Now, let

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

Then

$$(\psi^0, \varphi) \neq 0.$$

Then we have

$$\varphi^0 z^0(v) + \langle \varphi, \eta - F'(\bar{y}) \rangle - \langle F'(\bar{y})^* \varphi, z(\cdot; v) \rangle \geq 0, \quad \forall u \in U, \eta \in Q \tag{2.19}$$

Take $v = \bar{u}$, we obtain (1.9).
Next, we let $\eta = F(\bar{y})$ to get

$$\psi^0 z^0(v) - \langle F'(\bar{y})^* \varphi, z(\cdot; v) \rangle \leq 0 \ \forall\ v \in U. \tag{2.20}$$

Because $F'(\bar{y})^* \varphi \in W^{-1,p'}(\Omega)$, for the given ψ^0, there exists a unique solution $\psi \in W^{1,p'}(\Omega)$ of the adjoint Equation (1.10). Then, from (1.6), (2.16), and (2.2), we have

$$\begin{aligned}
0 &\geq \varphi^0 z^0(v) - \langle F'(\bar{y})^* \varphi, z(\cdot; v) \rangle \\
&= \psi^0 \int_\Omega f_y^0(x, \bar{y}, \bar{u}) z(x) \mathrm{d}x \\
&\quad + \int_\Omega \left[f^0(x, \bar{y}, v) - f^0(x, \bar{y}, \bar{u}) \right] \mathrm{d}x \\
&\quad + \langle -\nabla(\mu \nabla \bar{\psi}) - \nabla(\beta \bar{\psi}) + \sigma \bar{\psi} - \psi^0 f_y^0(x, \bar{y}, \bar{u}), z \rangle \\
&= \int_\Omega \{ \psi^0 \left[f^0(x, \bar{y}, v) - f_y^0(x, \bar{y}, \bar{u}) \right] \\
&\quad + \langle \psi, f(x, v) - f(x, \bar{u}) \rangle \} \mathrm{d}x \\
&= \int_\Omega \{ H(x, \bar{y}(x), v(x), \psi^0, \psi(x)) \\
&\quad - H(x, \bar{y}(x), \bar{u}(x), \psi^0, \psi(x)) \} \mathrm{d}x
\end{aligned} \tag{2.21}$$

There, (1.11) follows. Finally, by (1.10), if $(\psi^0, \psi) = 0$, then $F'(\bar{y})^* \varphi = 0$. Thus, in the case where

$$N\left(F'(\bar{y})^*\right) = \{0\},$$

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

[1] H. W. Lou and J. M. Yong, "Optimal Controls for Semilinear Elliptic Equations with Leaing Term Containing Controls," *SIAM Journal on Control and Optimization*, Vol. 48, No. 4, 2009, pp. 2366-2387. doi:10.1137/080740301

[2] J. T. Oden and J. N. Reddy, "Variational Methods in Theoretical Mechanics," Springer, Berlin and Heidelberg, 1983. doi:10.1007/978-3-642-68811-9

[3] L. Dede and A. Quarteroni, "Optimal Control and Numerical Adaptivity for Advection-Diffusion Equations," *Mathematical Modelling and Numerical Analysis*, Vol. 39, No. 2, 2005, pp. 1019-1040.

[4] N. N. Yan and Z. J. Zhou, "A Priori and a Posteriori Error Analysis of Edge Stabilization Galerkin Method for the Optimal Control Problem Governed by Convection-Dominated Diffusion Equation," *Journal of Computational and Applied Mathematics*, Vol. 223, No. 1, 2009, pp. 198-217. doi:10.1016/j.cam.2008.01.006

[5] R. Becker and B. Vexler, "Optimal Control of the Convection-Diffusion Equation Using Stabilized Finite Element Methods," *Numerische Mathematik*, Vol. 106, No. 3, 2007, pp. 349-367. doi:10.1007/s00211-007-0067-0

[6] S. Micheletti and S. Perotto, "An Anisotropic Mesh Adaptation Procedure for an Optimal Control Problem of the Advection-Diffusion-Reaction Equation," MOX-Report No. 15, 2008.

[7] S. S. Collis and M. Heinkenschloss, "Analysis of the Streamline Upwind/Petrov Galerkin Method Applied to the Solution of Optimal Control Problems," Technical Report 02-01, Department of Computational and Applied Mathematics, Rice University, Houston, 2002.

[8] X. Li and J. Yong, "Optimal Control Theory for Infinite-Dimensional Systems," Birkhäuser, Boston, 1995.

