Comments on `A discrete optimal control problem for descriptor systems'

Ravn, Hans V.

Published in:
IEEE Transactions on Automatic Control

Link to article, DOI:
10.1109/9.58518

Publication date:
1990

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
As we know,

\[ \text{In fact,} \]

exits and is continuous for

\[ \text{which is obviously positive and continuous on} (0, \infty) \times R \text{ and is } C^{1,1} \]
on (0, \infty) \times \{ \delta(x) \}.

Using Properties 5 and 6 in Proposition 1, one can easily verify

\[ \int_{-\infty}^{\infty} p(t, \delta) dx = 1 \text{ } \forall (t, x) \in (0, \infty) \times R. \]

In fact,

\[ \int_{-\infty}^{\infty} p(t, \delta) dx \]

\[ = \int_{-\infty}^{\infty} e^{x-y} G(t, x-y) H(1, t, |x-y|) dx \]

\[ + \int_{0}^{\infty} e^{-x-y} G(t, x+y) H(-1, t, |x-y|) dy \]

\[ = 1. \]

Next, we compute the mean \( u(t, x) \) by using (35).

As we know, \( u(t, x) \) solves the backward equation

\[ \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} = 0. \]

Computation gives

\[ u(t, x) = y + E [x - y] \]

\[ + \int_{-\infty}^{\infty} e^{x-y} G(t, x-y) dx \]

\[ = y + \int_{-\infty}^{\infty} e^{x-y} H(1, t, |x-y|) dx \]

\[ + \int_{0}^{\infty} e^{-x-y} H(-1, t, |x-y|) dy \]

\[ + (x - y). \]

It is easy to see \( u(t, x) \) is \( C^{1,1} \) on \( (0, \infty) \times R \) and \( \partial^2 u(t, x)/\partial x^2 \)
exists and is continuous for \( x \neq y. \)

The steady-state density function \( p_\ast(z) \) can be obtained by directly taking the limit

\[ p_\ast(z) = \lim_{t \to \infty} p(t, x | z) \]

\[ = \lim_{t \to \infty} e^{x-y} H(1, t, |x-y|) \]

\[ + \lim_{t \to \infty} e^{-x-y} H(-1, t, |x-y|) \]

\[ = e^{x-y} \left[ 1 - \Phi \left( \frac{-|x| - |y|}{\sqrt{2t}} \right) \right] \]

\[ = e^{-x-y} \left[ 1 - \Phi \left( \frac{|x| - |y|}{\sqrt{2t}} \right) \right] \]

\[ \text{where we have used the relation (6). And, of course, the invariant measure of (33) is } \mu(dz) = e^{-|z|} dz. \]

Before we conclude this example, let us make the following observation: let the diffusion \( \eta(t) \) be governed by

\[ d \eta(t) = -\text{sgn}(\eta(t) - y) dt + \text{sgn}[f(\eta(t))] dW(t) \]

\[ \text{where } f: R \to R \text{ is Lebesgue measurable.} \]

We claim that \( \eta(t) \) and \( \eta'(t) \), determined by (33), share the same transition probability density given in (35) and the same Fokker-Planck equation (34).

In fact, to see this, it is sufficient to notice that

\[ W(t) = \int_{0}^{t} \text{sgn}[f(\eta(s))] dW(s) \]

is another Brownian motion because \( W(t), \xi, t \geq 0 \) and \( W^2(t) - t, \xi, t \geq 0 \) are both martingales. Therefore, (36) can be rewritten as

\[ d \eta(t) = -\text{sgn}(\eta(t) + y) dt + dW(t), \]

i.e., \( \eta'(t) \) of (33) is a weak solution of (36).

References


Comments on “A Discrete Optimal Control Problem for Descriptor Systems”

HANS F. RAVN

Abstract — In a recent paper, necessary and sufficient optimality conditions are derived for a discrete-time optimal control problem, as well as other specific cases of implicit and explicit dynamic systems. We correct a mistake and demonstrate that there is not an “if and only if” correspondence between stationarity conditions and minimization of the Hamiltonian.

Manuscript received April 10, 1989.

The author is with the Institute of Mathematical Statistics and Operations Research, The Technical University of Denmark, DK-2800 Lyngby, Denmark.

IEEE Log Number 9036554.

I. Introduction
In the paper, the following problem was considered:

\[
\min \ G(x_N) + \sum_{k=0}^{N-1} L_k(x_k, u_k)
\]

(1a)

\[
x_{k+1} = f_k(x_k, u_k), \quad k = 0, \ldots, N-1
\]

(1b)

\[
qu_k(x_k, u_k) \geq 0, \quad k = 0, \ldots, N-1
\]

(1c)

\[
x_0 = x_0
\]

(1d)

where \( x_k \in R^n, u_k \in R^m, G: R^m \rightarrow R, L_k: R^n \rightarrow R, f_k: R^n \rightarrow R^n \), and as other specific cases of implicit and explicit dynamic systems.

In this note, we correct an error in the paper and extend the results by weakening the assumptions on constraint qualifications.

The approach taken in the paper, as well as here, is to derive optimality conditions for a specific case of a nonlinear programming problem. In this approach, a central element is the derivation of the Kuhn-Tucker conditions, and the identification of assumptions under which these conditions are necessary and/or sufficient, respectively, for optimality. This is supplemented with the control approach, where the Kuhn-Tucker stationarity conditions are supplemented with (or partially substituted by) minimization of the Hamiltonian with respect to the control \( u_k \).

II. Main Results
Let us introduce the following assumptions.

Assumption 1: \( L_k, f_k, q_k, k = 0, \ldots, N-1 \) and \( G \) are continuously differentiable with respect to all their variables.

Assumption 2: The Mangasarian-Fromowitz constraint qualification holds at the optimal points \( (x_k^*, u_k^*) \), \( k = 0, \ldots, N \), i.e., there holds that:

i) the gradients \( V^F(y) \) are linearly independent, and

ii) there exists a \( z \in R^{N(m+n)} \) such that

\[
\nabla F(x^*)z = 0
\]

\[
\nabla q_k(x^*)z > 0
\]

for all \( (i, k) \) for which \( q_k(x^*) = 0 \);

here

\[
y = (x_0, \ldots, x_N, u_0, \ldots, u_{N-1})^T,
\]

\[
F = (f_0, \ldots, f_N), \quad f_k = (f_{i,1}, \ldots, f_{i,n}),
\]

\[
q = (q_0, \ldots, q_N), \quad q_k = (q_{i,0}, \ldots, q_{i,n}).
\]

Remark 1: This assumption is weaker than the assumption of linear independence of \( \nabla F(x^*) \) and those \( \nabla q_k(x^*) \) for which \( q_k(x^*) = 0 \), which was used in the paper.

Assumption 3: At the optimal solution \( x_k^* \), \( L_k \) is convex, \( f_k \) is affine, and \( q_k \) is quasi-concave with respect to \( u_k \), \( k = 0, \ldots, N-1 \).

Assumption 4: \( L_k \) is pseudoconvex, \( f_k \) is quasi-linear (i.e., quasi-convex and quasi-concave) and \( q_k \) is quasi-concave at \( (x_k, u_k) \), \( k = 0, \ldots, N-1 \); and \( G \) is pseudoconvex at \( x_N \).

Theorem 1: If \((x, u)^* = (x_0^*, \ldots, x_N^*, u_0^*, \ldots, u_{N-1}^*)\) is optimal in (1), then under Assumptions 1 and 2, there exist vectors \( \lambda_k \in R^n \) and \( \mu_k \in R^n \) such that at \((x, u)^*\) there hold

\[
\lambda_k = \frac{\partial H_k}{\partial x_k} - \frac{\partial q_k}{\partial u_k} \mu_k, \quad k = 1, \ldots, N-1
\]

(2a)

\[
\lambda_N = \frac{\partial G}{\partial x_N}
\]

(2b)

\[
x_{k+1} = \frac{\partial H_k}{\partial x_k} \quad k = 0, \ldots, N-1
\]

(2c)

\[
x_0 = x_0
\]

(2d)

\[
\frac{\partial H_k}{\partial u_k} - \frac{\partial q_k}{\partial u_k} \mu_k = 0, \quad k = 0, \ldots, N-1
\]

(2e)

where

\[
L_k(x_k, u_k, \lambda_{k+1}) = L_k(x_k, u_k) + \lambda_{k+1} f_k(x_k, u_k)
\]

(2g)

If Assumption 3 holds also, then for \( k = 0, \ldots, N-1 \), \( u_k^* \) is a solution to

\[
\min H_k(x_k^*, v, \lambda_{k+1})
\]

(3a)

\[
qu_k(x_k^*, u_k^*) \geq 0.
\]

(3b)

Proof: The first result is proved in [5]. For the second result observe, that for \( x_k = x_k^* \), the Hamiltonian \( H_k \) is convex \( (L_k \text{ convex, } f_k \text{ affine, and hence, also } L_k, f_k \text{ affine}) \) with respect to \( u_k \) and therefore, also pseudoconvex [1, p. 108] and \( q_k \) quasi-concave with respect to \( u_k \). Therefore, the conditions (2e)-(2g) are sufficient for optimality of \( u_k^* \) in (3) [1, pp. 147-148].

Remark 2: The assumption of convexity of \( L_k \) in the last part of Theorem 1 cannot be substituted by an assumption of pseudoconvexity of \( L_k \).

Remark 3: This result can also be obtained under the following weaker Assumption 4: \( f_k \) is continuously differentiable, \( f_k \) is quasi-linear (i.e., quasi-convex and quasi-concave) and \( q_k \) is quasi-concave at \( (x_k, u_k) \), \( k = 0, \ldots, N-1 \); and \( G \) is pseudoconvex at \( x_N \).

Theorem 2: Assume that Assumption 1 holds, and that there exist \( \lambda, \mu \) such that (2) holds at \((x, u)^*\). If Assumption 4 holds also, then \((x, u)^* \) is optimal in (1).

Proof: We first show that the criterion function (1a) is pseudoconvex. The key observation is that (1a) is additive (viz. the sum of \( L_k \), \( k = 0, \ldots, N-1 \), and \( G \)) since all terms in (1a) are continuously differentiable (1a) is continuously differentiable; therefore, the gradient is zero, if and only if any partial derivative is zero. If the partial derivative with respect to \( (x_k, u_k) \) is zero, then \( L_k \) attains a minimum since \( L_k \) is pseudoconvex, and similarly holds for \( G \). Since (1a) is additive, the attainment of a minimum in each term implies that (1a) attains a minimum. Therefore, (1a) is pseudoconvex. Now, the result is proved as in [1, pp. 147-148] by observing that (2c), (2f), and (2g) imply that \((x, u)^* \) is feasible in (1).

Remark 4: In Theorem 2, the stationarity condition (2e) cannot be substituted by the condition that \( u_k^* \) is optimal in (3). However, (2e) may be substituted by the condition that \( u_k^* \) is an optimal solution to

\[
\min H_k(x_k^*, v, \lambda_{k+1}) - \mu_k q_k(x_k^*, u_k^*)
\]

(4)

But this condition is actually stronger than (2e); since (4) is an unconstrained problem with a continuously differentiable criterion function, the optimal point in (4) is a stationary point [1, p. 125] and this implies that (2e) holds.

III. Discussion
We have given necessary and sufficient optimality conditions for a discrete-time optimal control problem. The conditions are derived from similar stationary conditions in nonlinear programming, and supplemented by conditions from the control approach, in which the Hamiltonian is minimized. It is shown that the distinction between convexity and pseudoconvexity is essential, and that the results from the two approaches thus differ, implying that there is not an "if and only if" correspondence between stationarity conditions and minimization of the Hamiltonian.
The discussion about the equivalence or nonequivalence between various versions of optimality conditions in connection with discrete-time optimal control is old (see [7]). The mathematical programming approach has been most extensively treated in [2]. Derivation of optimality conditions from the saddle-point theorem of mathematical programming was done in [8]. A discussion of the connection between mathematical programming and discrete-time optimal control was performed in [4].

In all the aforementioned references, the Hamiltonian was defined as in (2g). By a suitable generalization of the Hamiltonian it is possible to specify weaker assumptions under which the Hamiltonian is minimized (see, e.g., [3], [6], or [7]).

**References**

[7] [The revised version of Theorem 2 in the paper’ is given by the follow- ing theorem:]

**Theorem 2.2:** Consider the control problem (19). Let \( L_k \) be convex, and \( q_k \) be quasi-concave in \( x_k \) and \( u_k \), \( k = 1, \ldots, N - 1 \). If the sequence \((x_k, u_k), k = 1, \ldots, N\) is an optimal solution to the problem, then there exist vectors \( l_1, \ldots, l_N, \

\text{and discussions are made.}

**Consider a linear system whose characteristic polynomial depends on \( p \) physical parameters \( q_j \) with \( q_j \in \{q_j^-, q_j^+\}, j = 1, 2, \ldots, p \). Suppose that the characteristic polynomial is of the form

\[ p(s, q) = \sum_{j=0}^{p} a_j(q) s^j \]

and the coefficient perturbations are polytopic. Then the family of polynomials \( P \)

\[ P = \{p(s, q); q \in Q \cap R^p\} \]

This implies that \( D \) is the half plane described by \( \Re \sigma \leq \sigma_r \).

Let \( D \) be the union of a finite number \((\geq 1)\) of pathwise connected regions in the complex plane. Define the notation \( \phi \) as a continuous function in \( p \) and \( d \) such that \((2a)-(2f)\) hold.

**Theorem 2.2:** Consider the problem (19). Let \( L_k \) be convex, and \( q_k \) be quasi-concave in \( x_k \) and \( u_k \), \( k = 0, 1, \ldots, N - 1 \). If the sequence \((x_k, u_k), k = 1, \ldots, N\) is an optimal solution to the problem (19).

**The discussion of the equivalence or nonequivalence between various versions of optimality conditions in connection with discrete-time optimal control is old (see [7]). The mathematical programming approach has been most extensively treated in [2]. Derivation of optimality conditions from the saddle-point theorem of mathematical programming was done in [8]. A discussion of the connection between mathematical programming and discrete-time optimal control was performed in [4]. **

**In all the aforementioned references, the Hamiltonian was defined as in (2g). By a suitable generalization of the Hamiltonian it is possible to specify weaker assumptions under which the Hamiltonian is minimized (see, e.g., [3], [6], or [7]).**

**Authors’ Reply**

**JING-YUE LIN AND ZI-HOU YANG**

The authors would like to thank Prof. Ravn for his comments on the paper. While we appreciated the comments, we wish to give a revised version of Theorem 2 in the paper in the context of the rest of this response, to achieve a balance of emphasis on the control problem for descriptor systems which has not been adequately explored in the literature.

The revised version of Theorem 2 in the paper is given by the following theorems without proofs which can be given by a slight modification of those in the paper, according to the correction given by Prof. Ravn.

**Theorem 2.1:** Consider the control problem (19). Let \( L_k \) be convex, and \( q_k \) be quasi-concave in \( x_k \) and \( u_k \), \( k = 0, 1, \ldots, N - 1 \). If the sequence \((x_k, u_k), k = 1, \ldots, N\) is an optimal solution to the problem, then there exist vectors \( l_1, \ldots, l_N, u_0, \ldots, u_{N-1} \) such that \((20a)-(20f)\) hold.

**Theorem 2.2:** Consider the problem (19). Suppose the necessary conditions in Theorem 2.1 hold. If \( G \) is pseudoconvex in \( x_k, L_k \) is pseudoconvex and \( q_k \) is quasi-concave in \( x_k \) and \( u_k \), \( k = 0, 1, \ldots, N - 1 \), then the sequence \((x_k, u_k), k = 1, \ldots, N\) is an optimal solution to the problem (19).

**The discussion of the equivalence or nonequivalence between various versions of optimality conditions in connection with discrete-time optimal control is old (see [7]). The mathematical programming approach has been most extensively treated in [2]. Derivation of optimality conditions from the saddle-point theorem of mathematical programming was done in [8]. A discussion of the connection between mathematical programming and discrete-time optimal control was performed in [4].**

**In all the aforementioned references, the Hamiltonian was defined as in (2g). By a suitable generalization of the Hamiltonian it is possible to specify weaker assumptions under which the Hamiltonian is minimized (see, e.g., [3], [6], or [7]).**

**Comments on “A Generalization of Kharitonov’s Concept for Robust Stability Problems with Linearly Dependent Coefficient Perturbations”**

**YAU-TARNG JUANG**

**Abstract—** It is shown by a counterexample that the main theorem in the above paper may lead to an erroneous D-stability conclusion for certain polynomials among the considered ones. Suggestions are presented and discussed.

**I. INTRODUCTION**

The Kharitonov stability theorem [1] has attracted much attention to the robust stability problem in the recent literature. Based on Kharitonov’s four-polynomial concept, a generalization theorem for robust D-stability assurance of polynomials with linearly dependent coefficient perturbations is presented in the paper. In this note, we give a counterexample to show that the main theorem in the paper may have a misleading result. Subsequently, suggestions and discussions are made.

Consider a linear system whose characteristic polynomial depends on \( p \) physical parameters \( q_j \) with \( q_j \in \{q_j^-, q_j^+\}, j = 1, 2, \ldots, p \). Suppose that the characteristic polynomial is of the form

\[ p(s, q) = \sum_{j=0}^{p} a_j(q) s^j \]

where \( q = [q_1, q_2, \ldots, q_p]^T \) and the coefficient perturbations are polytopic. Then the family of polynomials \( P \)

\[ P = \{p(s, q); q \in Q \cap R^p\} \]

This implies that \( D \) is the half plane described by \( \Re \sigma \leq \sigma_r \).

Let \( D \) be the union of a finite number \((\geq 1)\) of pathwise connected regions in the complex plane. Define the notation \( \phi \) as a continuous function in \( p \) and \( d \) such that \((2a)-(2f)\) hold.

**Theorem 1:** Consider the problem (19). Suppose the necessary conditions in Theorem 2.1 hold. If \( G \) is pseudoconvex in \( x_k, L_k \) is pseudoconvex and \( q_k \) is quasi-concave in \( x_k \) and \( u_k \), \( k = 0, 1, \ldots, N - 1 \), then the sequence \((x_k, u_k), k = 1, \ldots, N\) is an optimal solution to the problem (19).

**The discussion of the equivalence or nonequivalence between various versions of optimality conditions in connection with discrete-time optimal control is old (see [7]). The mathematical programming approach has been most extensively treated in [2]. Derivation of optimality conditions from the saddle-point theorem of mathematical programming was done in [8]. A discussion of the connection between mathematical programming and discrete-time optimal control was performed in [4].**

**In all the aforementioned references, the Hamiltonian was defined as in (2g). By a suitable generalization of the Hamiltonian it is possible to specify weaker assumptions under which the Hamiltonian is minimized (see, e.g., [3], [6], or [7]).**