Bulletin of the *Transilvania* University of Braşov • Vol 7(56), No. 1 - 2014
Series III: Mathematics, Informatics, Physics, 139-148

ON THE GRADIENT METHOD APPLIED TO OPTIMAL CONTROL PROBLEM

Ernest SCHEIBER[1]

Abstract

The purpose of this paper is to give direct proofs of some convergence results for the gradient and gradient projection methods applied to optimal control problem. The methods are considered in the continuous approach. In the context of optimal control problem, direct proofs are given to the results from the point of view of the teaching methods. The gradient projection method is studied in a modified variant, which is programmed in a simpler way in a Computer Algebra System (CAS) environment. For simple optimal control problems it is possible to use symbolic computations. An example is solved in *Mathematica* CAS.

2010 *Mathematics Subject Classification:* 49N05, 97N80, 49M15
Key words: optimal control problem, gradient method, gradient projection method.

1 Introduction

The purpose of this paper is to give direct proofs of some convergence results for the gradient and gradient projection methods applied to optimal control problem. The methods are considered in the continuous approach.

The used framework is similar to that used in [5], to prove a convergence result for the extragradient method applied to optimal control problem.

The theory of gradient and gradient projection methods are well known in abstract spaces [2], in finite dimensional spaces [3, 7, 8, 9] and even for optimal control problem [4], but here we are interested to give direct proofs of the results in the context of optimal control problem from the point of view of the teaching methods.

In [4], an adaptive precision algorithm for the gradient method is developed and justified. The algorithm is designed to be implemented by numerical methods. The adaptive precision algorithm supposes the use of low accuracy numerical integration while the computations are far from the solution and improve the accuracy as the solution is approached.

The gradient projection method [3, 7, 8, 9] is studied in a modified variant, which is more adequate to be used in a Computer Algebra System (CAS) environment.

[1]Faculty of Mathematics and Informatics, *Transilvania* University of Braşov, Romania, e-mail: scheiber@unitbv.ro

For simple optimal control problems it is possible to use symbolic computations. Such an example is solved in *Mathematica* CAS.

2 The gradient method

We consider the optimal control problem

$$\text{minimize } I(u) = \varphi(x(T)) \tag{1}$$

subject to the constraints

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) && (2)\\ x(0) &= x^0 && (3) \end{aligned}$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^q, A(t) \in M_{n,n}(\mathbb{R}), B(t) \in M_{n,q}(\mathbb{R})$ are real matrix with continuous elements and φ is a continuous differentiable function. T is fixed and the set of admissible controls $\mathcal{U}$ is the set of piecewise continuous functions defined on $[0,T]$.

Let us denote by φ_x the derivate of φ and by M^* the transpose of a matrix M.

Let $X(t) \in M_{n,n}(\mathbb{R})$ be a fundamental solution matrix for the linear differential system $\dot{x}(t) = A(t)x(t)$ (i.e. $\dot{X}(t) = A(t)X(t)$ and $X(t)$ is not singular) and $\Phi(t,s) = X(t)X^{-1}(s)$.

The solution of the initial value problem (2)-(3) will be

$$x^u(t) = \Phi(t,0)x^0 + \int_0^t \Phi(t,s)B(s)u(s)\mathrm{d}s. \tag{4}$$

The matrix $[\Phi^*(t)]^{-1}$ is the fundamental solution matrix for the adjoin system

$$\dot{p}(t) = -A^*(t)p(t). \tag{5}$$

If

$$p(T) = p^T \tag{6}$$

then the solution of the problem (5)-(6), the costate, will be, [6],

$$p(t) = \Phi^*(T,t)p^T. \tag{7}$$

We introduce the controllability Gramian

$$W = \int_0^T \Phi(T,t)B(t)B^*(t)\Phi^*(T,t)\mathrm{d}t. \tag{8}$$

The following properties are known, [6]:

Theorem 2.1 *(i) W is symmetric and non-negative defined matrix;*
(ii) If W is invertible and $\hat{u}(t) = B^*(t)\Phi^*(T,t)\bar{u}$ with $\bar{u} = W^{-1}(x^T - \Phi(T,0)x^0)$ then $x^{\hat{u}}(t)$ defined by (4) verifies the boundary conditions $x^{\hat{u}}(0) = x^0$ and $x^{\hat{u}}(T) = x^T$.

(iii) Among all controls $u \in \mathcal{U}$ transferring x^0 to x^T in time T, the control $\hat{u}(t)$ minimizes the functional $\int_0^T \|u(t)\|_2^2 \mathrm{d}t$, moreover,

$$\int_0^T \|\hat{u}(t)\|_2^2 \mathrm{d}t = \int_0^T \|u(t)\|_2^2 \mathrm{d}t + \int_0^T \|\hat{u}(t) - u(t)\|_2^2 \mathrm{d}t.$$

A simple procedure to solve the optimal control problem (1)-(3) results: Compute the local minimizer points of $\varphi(x)$ and for any such point x^T solves the initial value problem

$$\begin{array}{l} \dot{x}(t) = A(t)x(t) + B(t)B^*(t)\Phi^*(T,t)W^{-1}(x^T - \Phi(T,0)x^0) \\ x(0) = x^0. \end{array}$$

Example 2.1

$$\text{minimize } x_1^4(T) - 2x_1^3(T) + x_1^2(T) + x_2^2(T)$$

subject to

$$\begin{array}{ll} \dot{x}_1(t) = x_2(t) & \qquad x_1(0) = x_1^0 \\ \dot{x}_2(t) = u(t) & \qquad x_2(0) = x_2^0 \end{array}$$

For this example

$$\Phi(t,s) = \begin{pmatrix} 1 & t-s \\ 0 & 1 \end{pmatrix}, \quad W = \begin{pmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{pmatrix}$$

and the results are

$$\begin{array}{rcl} \hat{u}(t) & = & \left(\frac{12(T-t)}{T^3} - \frac{6}{T^2}\right)(x_1^T - x_1^0 - Tx_2^0) + \left(\frac{4}{T} - \frac{6(T-t)}{T^2}\right)(x_2^T - x_2^0) \\ x_1(t) & = & x_1^0 + tx_2^0 + \frac{t^3(2x_1^0 + Tx_2^0 - 2x_1^T + Tx_2^T)}{T^3} + \frac{t^2\left(-3Tx_1^0 - 2T^2x_2^0 + 3Tx_1^T - T^2x_2^T\right)}{T^3} \\ x_2(t) & = & x_2^0 + \frac{t^2(6x_1^0 + 3Tx_2^0 - 6x_1^T + 3Tx_2^T)}{T^3} + \frac{t\left(-6Tx_1^0 - 4T^2x_2^0 + 6Tx_1^T - 2T^2x_2^T\right)}{T^3} \end{array}$$

Note that W is a strict posive defined matrix.
The function $\varphi(x) = x_1^4 - 2x_1^3 + x_1^2 + x_2^2$ has two minimum points $x^T = (0,0)^*$ and $x^T = (1,0)^*$. For these endpoints we obtain

Endpoint	$u(t)$	$x_1(t)$	$x_2(t)$
$(0,0)^*$	$-16+30t$	$2+t-8t^2+5t^3$	$1-16t+15t^2$
$(1,0)^*$	$-10+18t$	$2+t-5t^2+3t^3$	$1-10t+9t^2$

Especially, for numerical computations the gradient method is a more practical approach. In order to state the relations to be used, we need the expression of the gradient of the functional $I(u)$, [1].

Theorem 2.2 *The gradient of the functional $I(u)$ is*

$$\nabla I(u)(t) = -B^*(t)p^u(t) = -B^*(t)\Phi^*(T,t)p^{T,u} \tag{9}$$

with

$$p^u(T) = p^{T,u} = -\varphi_x(x^u(T)). \tag{10}$$

Proof. Due to the linearity of (2), for any $u, \delta u \in \mathcal{U}$ and $\lambda \in \mathbb{R}$ the state corresponding to the control $u + \lambda\delta u$ is $x^u + \lambda x^{\delta u}$ with $\dot{x}^{\delta u}(t) = A(t)x^{\delta u}(t) + B(t)\delta u(t)$ and $\delta x(0) = 0$. We compute

$$\lim_{\lambda\to 0}\frac{1}{\lambda}(I(u+\lambda\delta u)-I(u)) = \lim_{\lambda\to 0}\frac{1}{\lambda}(\varphi(x^u(T)+\lambda\delta x(T))-\varphi(x^u(T)) =< \varphi_x(x^u(T)), \delta x(T) > .$$

Using (10) and doing a standard computation, we have

$$< \varphi_x(x^u(T)), \delta x^u(T) >= - < p^u(T), \delta x^u(T) >=$$

$$= -\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} < p^u(t), \delta x^u(t) > \mathrm{d}t = -\int_0^T < B^*(t)p^u(t), \delta u(t) > \mathrm{d}t.$$

Considering (7), we obtain

$$< \varphi_x(x^u(T)), \delta x^u(T) >= -\int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, \delta u(t) > \mathrm{d}t.$$

Because the functional $\delta u \mapsto \int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, \delta u(t) > \mathrm{d}t$ is continuous, the expression of the gradient of $I(u)$ is given by (9). ∎

We suppose that the gradient of φ satisfies the Lipschitz condition

$$\|\varphi_x(y) - \varphi_x(x)\|_2 \le L\|y - x\|_2, \quad \forall x, y \in \mathbb{R}^n. \tag{11}$$

The following inequality is the starting point for the gradient method, [8, 9].

Theorem 2.3 *For any* $v, u \in \mathcal{U}$,

$$I(v)-I(u) \le \frac{L}{2}\|x^v(T)-x^u(T)\|_2^2 - \int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, v(t)-u(t) > \mathrm{d}t. \tag{12}$$

Proof. Using the technique from [8, 9], we have

$$I(v) - I(u) = \varphi(x^v(T)) - \varphi(x^u(T)) =$$

$$= \int_0^1 < \varphi_x(x^u(T) + s(x^v(T) - x^u(T))) - \varphi_x(x^u(T)), x^v(T) - x^u(T) > \mathrm{d}s +$$

$$+ < \varphi_x(x^u(T)), x^v(T) - x^u(T) > .$$

After applying the Cauchy inequality and (11) to the first term, the above inequality becomes

$$I(v) - I(u) \le \frac{L}{2}\|x^v(T) - x^u(T)\|_2^2 + < \varphi_x(x^u(T)), x^v(T) - x^u(T) > .$$

As in the proof of the above theorem,

$$< \varphi_x(x^u(T)), x^v(T) - x^u(T) >= -\int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, v(t) - u(t) > \mathrm{d}t. \quad ∎$$

We fix the control function $v(t)$ as

$$v(t) = u(t) - \alpha \nabla I(u)(t) = u(t) + \alpha B^*(t)\Phi^*(T,t)p^{T,u}, \qquad \alpha > 0. \tag{13}$$

Then, applying (4)

$$x^v(T) - x^u(T) = \alpha \int_0^T \Phi(T,t)B(t)B^*(t)\Phi^*(T,t)p^{T,u}\mathrm{d}t = \alpha W p^{T,u},$$

and

$$\int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, v(t) - u(t) > \mathrm{d}t = \alpha < W p^{T,u}, p^{T,u} > .$$

Using the above equalities, the inequality (12) is

$$I(v) - I(u) \leq \frac{L}{2}\alpha^2 \|W p^{T,u}\|_2^2 - \alpha < W p^{T,u}, p^{T,u} > . \tag{14}$$

If the matrix W is strictly positive defined then there exists $\omega > 0$ such that $< Wx, x > \geq \omega \|x\|_2^2,\ \forall x \in \mathbb{R}^q$. The inequality (14) becomes

$$I(v) - I(u) \leq (\frac{L}{2}\alpha^2 \|W\|_2^2 - \alpha\omega)\|p^{T,u}\|_2^2.$$

If $0 < \alpha < \frac{2\omega}{L\|W\|_2^2}$ then $I(v) - I(u) < 0$.

The gradient method consists in the construction of the sequence of control functions

$$\begin{aligned} u^{(0)} &\in \mathcal{U} \\ u^{(k+1)}(t) &= u^{(k)}(t) + \alpha_k B^*(t)p^{(k)}(t), \quad t \in [0,T], \qquad (15) \\ & \quad 0 < \alpha_k < \frac{2\omega}{L\|W\|_2^2}, \qquad \forall\ k \in \mathbb{N}. \qquad (16) \end{aligned}$$

To compute $u^{(k+1)}(t)$ two other functions are required: $x^{(k)}(t)$ and $p^{(k)}(t)$, the solutions of the initial value problems (2)-(3), with $u(t) = u^{(k)}(t)$, and respectively, (5)-(11), with $p^{T,u} = -\varphi_x(x^{(k)}(T))$.

Theorem 2.4 *If the gradient of $\varphi(x)$ satisfies the Lipschitz condition (11), the controllability matrix is strictly positive defined and the sequence $(u^{(k)})_{k\in\mathbb{N}}$ is defined by (15)-(16) then $(I(u^{(k)}))_{k\in\mathbb{N}}$ is a decreasing sequence.*

Theorem 2.5 *If $\varphi(x)$ is bounded below then the sequence $(I(u^{(k)}))_{k\in\mathbb{N}}$ is convergent.*

Requiring a stronger constraint to the parameters α_k it results that:

Theorem 2.6 *If $0 < \delta < \frac{\omega}{L\|W\|_2^2}$ and $\alpha_k \in (\delta, \frac{2\omega}{L\|W\|_2^2} - \delta),\ \forall k \in \mathbb{N}$, then the sequence $(p^{u^{(k)}}(t))_{k\in\mathbb{N}}$ converges uniformly to 0.*

Proof. From (14) it results that

$$I(u^{(k+1)}) - I(u^{(k)}) \leq \left(\frac{L}{2}\delta^2 \|W\|_2^2 - \delta\omega\right) \|p^{T,u^{(k)}}\|_2^2$$

or

$$\|p^{T,u^{(k)}}\|_2^2 \leq \frac{I(u^{(k)}) - I(u^{(k+1)})}{\delta\omega - \frac{L}{2}\delta^2 \|W\|_2^2}. \tag{17}$$

The right hand side of (17) converges to 0 and therefore $\lim_{k\to\infty} p^{T,u^{(k)}} = 0$. Using (7) we deduce that the sequence $(p^{u^{(k)}}(t))_{k\in\mathbb{N}}$ converges uniformly to 0. ■

For some simple problems, we may avoid numerical computations in favor of the symbolic computation. To solver the above example (2.1), using the gradient method, the *Mathematica* codes are

```
T = 1;
eps = 0.01;
B = {{0}, {1}};
u[t_] := 0
F[p_, q_] := p^4 - 2 p^3 + p^2 + q^2

Step[u_] :=
 Module[{s, cs, x1, x2, p1, p2, state, X1, X2, costate, P1, P2},
  state := {x1'[t] - x2[t] == 0, x2'[t] - u[t] == 0, x1[0] == 2,
    x2[0] == 1};
  s = DSolve[state, {x1[t], x2[t]}, t];
  X1[x_] := Last[s[[1, 1]]] /. t -> x;
  X2[x_] := Last[s[[1, 2]]] /. t -> x;
  costate := {p1'[t] == 0, p2'[t] + p1[t] == 0,
    p1[T] == -D[F[x, y], x] /. {x -> X1[T], y -> X2[T]},
    p2[T] == -D[F[x, y], y] /. {x -> X1[T], y -> X2[T]}};
  cs = DSolve[costate, {p1[t], p2[t]}, t];
  P1[x_] := Last[cs[[1, 1]]] /. t -> x;
  P2[x_] := Last[cs[[1, 2]]] /. t -> x;
  {u[t] + eps  Last[Last[Transpose[B].{{P1[t]}, {P2[t]}}]],
   N[F[X1[T], X2[T]]], X1[t], X2[t]}]

Grad[u_, n_] :=
 Module[{uu, i, s},
  For[i = 0, i < n, i++, uu = Step[u]; Clear[u];
  u:=Function[t,uu[[1]]]; If[IntegerQ[i/100], Print[uu[[2]]]]];
  Step[u]]

gg = Grad[u, 2000];
x1 = Expand[N[gg[[3]]]]
2. + 1. t - 4.90559 t^2 + 2.93193 t^3
x2 = Expand[N[gg[[4]]]]
1. - 9.81118 t + 8.7958 t^2
uu = Expand[N[gg[[1]]]]
-9.81144 + 17.5922 t
```

In this case the endpoint is $(1,0)^*$ and the cost functional is 0.00128824. The values for α_k are $\alpha_k := eps = 0.01,\ \forall k \in \mathbb{N}$. For `u[t_]:=20t` the endpoint will be $(0,0)^*$.

3 The gradient projection method

Let U be a convex and close subset of $\mathbb{R}^q$ and the constraint

$$u(t) \in U, \qquad \forall\, t \in [0,T]. \tag{18}$$

In this case, a control function $u(t)$ will be admissible if it is piecewise continuous and satisfies the constraint (18).

If $u \in \mathbb{R}^q$ and $v \in U$ is the projection of u into U, $v = \mathrm{Pr}_U(u)$, then

$$\|v-u\|_2 = \min_{w\in U} \|w-u\|_2 \quad \Leftrightarrow \quad < v-u, v-w > \leq 0,\ \forall w \in U \tag{19}$$

and consequently

$$\|\mathrm{Pr}_U(x) - \mathrm{Pr}_U(y)\|_2^2 \leq < x-y, \mathrm{Pr}_U(x) - \mathrm{Pr}_U(y) >, \qquad \forall\ x, y \in \mathbb{R}^q. \tag{20}$$

The simplest form of the gradient projection method is defined by

$$u^{(k+1)}(t) = \mathrm{Pr}_U(u^{(k)}(t) - \alpha_k \nabla I(u^{(k)})(t),\ t \in [0,T],\ k \in \mathbb{N}.$$

Our implementation in *Mathematica* uses the variant

$$\begin{aligned} u^{(0)} &= \tilde{u}^{(0)} \in \mathcal{U}, \\ \tilde{u}^{(k+1)}(t) &= \tilde{u}^{(k)}(t) - \alpha_k \nabla I(u^{(k)})(t), \quad t \in [0,T],\ k \in \mathbb{N}, \qquad (21) \\ u^{(k+1)}(t) &= \mathrm{Pr}_U(\tilde{u}^{(k)}(t)). \qquad (22) \end{aligned}$$

In order to apply (12) we must evaluate the two terms of its right side.

Given $\tilde{u} : [0,T] \to \mathbb{R}^q$ piecewise continuous and $\alpha > 0$, let be

$$\begin{aligned} u(t) &= \mathrm{Pr}_U(\tilde{u}(t)) \\ \tilde{v}(t) &= \tilde{u}(t) - \alpha \nabla I(u)(t) = \tilde{u}(t) + \alpha B^*(t)\Phi^*(T,t)p^{T,u} \\ v(t) &= \mathrm{Pr}_U(\tilde{v}(t)), \qquad t \in [0,T]. \end{aligned}$$

Denoting $\Psi(t,s) = B^*(s)\Phi^*(T,s)\Phi(T,t)B(t) \in M_{q,q}(\mathbb{R})$ and $C_\Psi =$ $= (\int_0^T\int_0^T \|\Psi(t,s)\|_2^2 \mathrm{d}t\mathrm{d}s)^{\frac{1}{2}}$, we obtain successively

$$\|x^v(T) - x^u(T)\|_2^2 = < x^v(T) - x^u(T), x^v(T) - x^u(T) > =$$

$$= < \int_0^T \Phi(T,t)B(t)(v(t)-u(t))\mathrm{d}t, \int_0^T \Phi(T,s)B(s)(v(s)-u(s))\mathrm{d}s > =$$

$$= \int_0^T\int_0^T < \Psi(t,s)(v(t)-u(t), v(s)-u(s) > \mathrm{d}t\mathrm{d}s.$$

Applying successively the Cauchy's inequalities, we obtain

$$\|x^v(T) - x^u(T)\|_2^2 \leq \int_0^T\int_0^T \|\Psi(t,s)(v(t)-u(t)\|_2 \|v(s)-u(s)\|_2 \mathrm{d}t\mathrm{d}s \leq$$

$$\leq \int_0^T\int_0^T \|\Psi(t,s)\|_2 \|v(t)-u(t)\|_2 \|v(s)-u(s)\|_2 \mathrm{d}t\mathrm{d}s =$$

$$= \int_0^T \|v(t)-u(t)\|_2 \left(\int_0^T \|\Psi(t,s)\|_2 \|v(s)-u(s)\|_2 \mathrm{d}s \right) \mathrm{d}t.$$

The internal integral is increased by

$$\int_0^T \|\Psi(t,s)\|_2\|v(s)-u(s)\|_2\mathrm{d}s \le \left(\int_0^T \|\Psi(t,s)\|_2^2\mathrm{d}s\right)^{\frac{1}{2}}\left(\int_0^T \|v(s)-u(s)\|_2^2\mathrm{d}s\right)^{\frac{1}{2}}.$$

Thus

$$\|x^v(T)-x^u(T)\|_2^2 \le \left(\int_0^T \|v(s)-u(s)\|_2^2\mathrm{d}s\right)^{\frac{1}{2}}\left(\int_0^T \|\Psi(t,s)\|_2^2\mathrm{d}s\right)^{\frac{1}{2}} \times$$

$$\times \int_0^T \|v(t)-u(t)\|_2\mathrm{d}t \le C_\Psi \int_0^T \|v(t)-u(t)\|_2^2\mathrm{d}t. \tag{23}$$

The second term of (12) is

$$\int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, v(t)-u(t) > \mathrm{d}t = \frac{1}{\alpha}\int_0^T < \tilde{v}(t)-\tilde{u}(t), v(t)-u(t) > \mathrm{d}t.$$

Using (20) it follows that

$$\int_0^T < B^*(t)\Phi^*(T,t)p^{T,u}, v(t)-u(t) > \mathrm{d}t \ge \frac{1}{\alpha}\int_0^T \|v(t)-u(t)\|_2^2\mathrm{d}t.$$

Finally, from (12), it results

$$I(v)-I(u) \le (\frac{L}{2}C_\psi - \frac{1}{\alpha})\int_0^T \|v(t)-u(t)\|_2^2\mathrm{d}t.$$

Theorem 3.1 *If the gradient of $\varphi(x)$ satisfies the Lipschitz condition (11) and the sequence $(u^{(k)})_{k\in\mathbb{N}}$ is defined by (21)-(22), with $(\alpha_k \in (0, \frac{2}{LC_\Psi})$, then $(I(u^{(k)}))_{k\in\mathbb{N}}$ is a decreasing sequence.*

Additionally, if $\varphi(x)$ is bounded below then the sequence $(I(u^{(k)}))_{k\in\mathbb{N}}$ is convergent.

The *Mathematica* codes to solve the problem from Example 2.1, with the additional constraint $|u(t)| \le 1$ are

```
Pr[u_, t_] := If[u[t] < -1, -1, If[u[t] > 1, 1, u[t]]]
eps = 0.01;
T = 1;
B = {{0}, {1}};
u[t_] := 0
F[p_, q_] := p^4 - 2 p^3 + p^2 + q^2

Step[u_] :=
 Module[{s, cs, x1, x2, p1, p2, state, X1, X2, costate, P1, P2},
  state := {x1'[t] - x2[t] == 0, x2'[t] - Pr[u, t] == 0, x1[0] == 2,
    x2[0] == 1};
  s = DSolve[state, {x1[t], x2[t]}, t];
  X1[x_] := Last[s[[1, 1]]] /. t -> x;
  X2[x_] := Last[s[[1, 2]]] /. t -> x;
  costate := {p1'[t] == 0, p2'[t] + p1[t] == 0,
    p1[T] == -D[F[x, y], x] /. {x -> X1[T], y -> X2[T]},
```

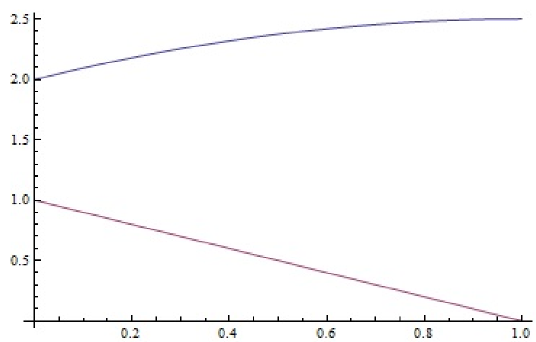
```
    p2[T] == -D[F[x, y], y] /. {x -> X1[T], y -> X2[T]}};
  cs = DSolve[costate, {p1[t], p2[t]}, t];
  P1[x_] := Last[cs[[1, 1]]] /. t -> x;
  P2[x_] := Last[cs[[1, 2]]] /. t -> x;
  {u[t] + eps  Last[Last[Transpose[B].{{P1[t]}, {P2[t]}}]],
   N[F[X1[T], X2[T]]], X1[t], X2[t]}]

GradProj[u_, n_] :=
 Module[{uu, i, s},
  For[i = 0, i < n, i++, uu = Step[u]; Clear[u];
  u:=Function[t,uu[[1]]]; If[IntegerQ[i/100], Print[uu[[2]]]]];
  uu = Step[u]]
```

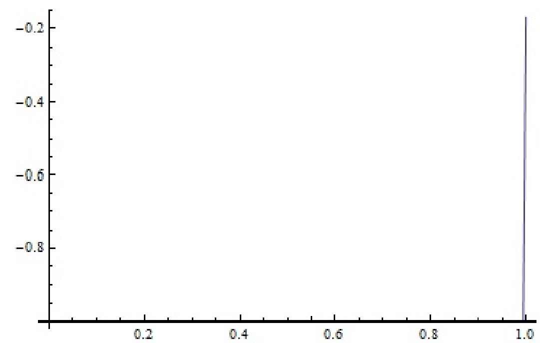
After 500 iteration the minimum obtained value is 14.0627 and the control function is

$$u[t] := \begin{cases} -1 & t < 0.994488 \\ 1 & t > 1.00773 \\ -151.153 + 150.985t & \text{True} \end{cases}$$

The plot of the state functions and the control function are



The state functions



The control function

It may be verified that $u(t) = -1$ is a solution of this optimal control problem. The corresponding states are $x_1(t) = 2 + t - \frac{t^2}{2},\ x_2(t) = 1 - t$ while the costates are $p_1(t) = -30,\ p_2(t) = 30(t-1)$.

Any variation $\delta u(t)$ of the control $u(t) = -1$ satisfies the constraint $\delta u(t) \in (0, 2)$. As usual, the corresponding state is denoted $\delta x(t) = (\delta x_1(t), \delta x_2(t))$.

The following equality holds

$$I(u+\delta u) = I(u) + < \nabla I(u), \delta u > + \int_0^1 (1-s) < \mathrm{d}\varphi^2(x^u(T) + s\delta x(T))\delta x(T), \delta x(T) > \mathrm{d}s =$$

$$= I(u) - \int_0^T < B^*(t)p^u(t), \delta u(t) > \mathrm{d}t +$$

$$+ \int_0^1 (1-s) < \begin{pmatrix} \frac{\partial^2\varphi}{\partial x_1^2}(x^u(T) + s\delta x(T)) & \frac{\partial^2\varphi}{\partial x_2 \partial x_1}(x^u(T) + s\delta x(T)) \\ \frac{\partial^2\varphi}{\partial x_1 \partial x_2}(x^u(T) + s\delta x(T)) & \frac{\partial^2\varphi}{\partial x_2^2}(x^u(T) + s\delta x(T)) \end{pmatrix} \begin{pmatrix} \delta x_1(T) \\ \delta x_2(T) \end{pmatrix}, \begin{pmatrix} \delta x_1(T) \\ \delta x_2(T) \end{pmatrix} > \mathrm{d}s.$$

Because $x^u(T) = (\frac{5}{2}, 0)$, the results of the computations are

$$I(u) = \frac{225}{16} = 14.0625;$$

$$-\int_0^T < B^*(t)p^u(t), \delta u(t) > \mathrm{d}t = 30\int_0^1 (1-t)\delta u(t)\mathrm{d}t \geq 0$$

and the last term is

$$\int_0^1 (1-s)\left((47+48s\delta x_1(T)+12s^2\delta x_1^2(T))\delta x_1^2(T)+\delta x_2^2(T)\right)\mathrm{d}s =$$

$$= \left(\frac{47}{2}+8\delta x_1(T)+\delta x_1^2(T)\right)\delta x_1^2(T)+\delta x_2^2(T) \geq 0.$$

The last two inequalities prove the assertion.

References

[1] Barbu V., *Mathematical methods in optimization of differential systems.* Ed. Acad. R.S.R., Bucureşti, 1989 (Romanian).

[2] Ilioi C., *Optimization problems and algorithms for approximate solutions.* Ed. Acad. R.S.R., Bucureşti, 1980 (Romanian).

[3] Kelley C. T., *Iterative Methods for Optimization.* SIAM, Philadelphia, 1999.

[4] Klessig R., Polak E., *An Adaptive Precision Gradient Method for Optimal Control.* SIAMJ. Control, **11**, No. 1,80-93, 1973.

[5] Khoroshilova V.E., *Extragradient-type method for optimal control problem with linear constraints and convex objective function.* Optim Lett., **7**, 1193-1214, 2013.

[6] Zabczyk J., *Mathematical Control Theory: An Introduction.* Ed. Birkhäuser, Boston, 1992.

[7] Васильев Ф. П., Численные методы решения экстремальных задач. Наука, Москва, 1980.

[8] Карманов В. Г., Математическое программирование. Наука, Москва, 1975.

[9] Пшеничный Б. Н., Данилин Ю. М., Численные методы в экстремальных задачах. Наука, Москва, 1975.