Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 6, n. 2, 2018.

Trabalho apresentado no XXXVIII CNMAC, Campinas - SP, 2018.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

---

O Princípio do Máximo para Problemas Não-Diferenciáveis de Controle Ótimo Discreto Multiobjetivos

Camila Isoton[1]
Universidade Positivo, UP, Curitiba, PR
Lucelina Batista Santos[2]
Departamento de Matemática, UFPR, Curitiba, PR

Resumo. Neste trabalho obtemos uma versão do Princípio do Máximo para uma certa classe de problemas de controle ótimo discreto multiobjetivos não diferenciáveis. Nosso resultado aqui estabelecido generaliza os obtidos por Shvartsman [11] e Mordukhovich e Shvartsman [8] para problemas multiobjetivos.

Palavras-chave. Problemas de Controle Ótimo Discretos, Problemas Multiobjetivos, Não Suavidade

1 Introdução e Formulação do Problema

Vários problemas em planejamento econômico, tecnologia, organização de produção e investigação de operações são descritos por sistemas de equações de diferenças [2, 10]. Além disso, os sistemas discretos são encontrados na resolução de problemas de controle ótimo contínuos por meio de métodos numéricos [10]. A programação dinâmica [2] é um método universal para investigar sistemas de controle ótimo discretos, apesar disso existem algumas desvantagens sobre outros métodos e a aplicação deste modelo em problemas específicos pode ser menos efetiva. Isto é o porquê na maioria dos casos (especialmente nos sistemas discretos de grande porte), o uso das condições de otimalidade se torna preferível ao da programação dinâmica.

A Teoria de Controle Ótimo Contínuo tem como um resultado central o Princípio do Máximo provado por Pontryagin *et al.* [9] e dada a sua significância o que se busca é um resultado análogo em tempo discreto. Como bem sabemos o princípio do máximo discreto não é válido em geral sem hipóteses de convexidade impostas *a priori* no sistema de controle. Depois disso, muitos autores tentaram provar o princípio do máximo em uma versão mais fraca (máximo local, estado de estacionariedade) e juntamente com condições de otimalidade de ordens maiores. Levando em consideração os estudos acima mencionados e os resultados dos artigos [1–3, 5–7], dentre outros, podemos dizer que a teoria das condições de otimalidade em sistemas discretos continua sendo um tema de pesquisa atual.

[1]isoton.camila@gmail.com
[2]lucelina@ufpr.br

Nosso interesse também está nos problemas em programação multiobjetivo, pois estes surgiram no sentido de contemplar situações mais realistas, nas quais se estudam vários objetivos, em geral conflitantes, que concorrem para uma determinada solução.

Por outro lado temos que o papel da Análise não-diferenciável é de fundamental importância na Teoria de Otimização e isto se deve a pelo menos dois motivos fundamentais: Por um lado, em muitas aplicações a hipótese de diferenciabilidade pode ser muito restritiva e, por outro lado, existem inúmeros mecanismos naturais que geram "não suavidades", mesmo quando se parte de problemas diferenciáveis. Isto ocorre, por exemplo, na Teoria de Dualidade, Análise de Estabilidade, Técnicas de Decomposição, Métodos de Penalização, dentre outros. Veja [4] e suas referências.

Daí, nosso interesse no estudo de problemas de controle ótimo discreto não diferenciáveis. Nessa direção, estendemos o teorema do Princípio do Máximo escalar provado por Schavartsman em [11] a problemas de controle ótimo discreto multiobjetivos. Mais precisamente, trataremos do seguinte problema de controle ótimo discreto multiobjetivo não suave:

$$\begin{array}{l} \text{Minimizar } (\Phi_0(x_1,\cdots,x_N),\cdots,\Phi_{r_0}(x_1,\cdots,x_N)) \\ \text{sujeito a:} \\ x_{i+1}=\varphi_i(x_i,u_i),\ i=0,\cdots,N-1 \\ x_0=\mathbf{x}_0 \\ \Phi_j(x_1,\cdots,x_N)\le 0,\ j=r_0+1,\cdots,r_0+r_1 \\ \Phi_j(x_1,\cdots,x_N)=0,\ j=r_0+r_1+1,\cdots,r_0+r_1+r_2 \\ u_i\in U_i,\ i=0,\cdots,N-1 \end{array} \tag{PND}$$

onde $\varphi : [0, N-1]^3 \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\Phi_j : \mathbb{R}^{n(N)} \to \mathbb{R}$, $j = 0, \cdots, r_0 + r_1 + r_2$ são funções dadas e $U_i \subset \mathbb{R}^m, i \in [0, N-1]$, são subconjuntos não vazios de $\mathbb{R}^m$.

Neste caso, a otimalidade será entendida no sentido da eficiência fraca. Mais precisamente, um processo admissível para (PND), $(\hat{x}, \hat{u})$ é um processo Pareto fraco se não existe (x, u) admissível para (PND) tal que $\Phi_j(x_1, \cdots, x_N) < \Phi_j(\hat{x}_1, \cdots, \hat{x}_N)$, para todo $j = 0, \cdots, r_0$.

O texto tem a seguinte estrutura: Na Seção 2, apresentamos alguns resultados básicos de Análise Não Suave e na Seção 3 demonstraremos o Princípio do Máximo Discreto para (PND).

2 Resultados Preliminares

Neste trabalho, assumiremos que o espaço $\mathbb{R}^n$ está munido do produto interno usual o qual denotaremos $\langle \cdot, \cdot \rangle$ e a norma associada a este produto será denotada $||\cdot||$. Admitiremos que este espaço está munido da topologia induzida pela norma $||\cdot||$.

Sejam Ω um subconjunto aberto e não vazio de $\mathbb{R}^n$, $\phi : \Omega \to \mathbb{R}$ uma função dada e fixados $x \in \Omega$ um ponto e $v \in \mathbb{R}^n$ uma direção. A **derivada direcional generalizada**

[3] $[0, N]$ é o intervalo discreto da reta real, com $N \geq 0$ o número de etapas (passos) realizados

DOI: 10.5540/03.2018.006.02.0307

(no sentido de Clarke) de ϕ em x, na direção v, é dada por

$$\phi^0(x;v) = \lim \sup_{\substack{y \to x \\ \lambda \to 0^+}} \frac{\phi(y+\lambda v) - \phi(y)}{\lambda}$$

e o **subdiferencial** (de Clarke) de ϕ em $x \in \Omega$ é o conjunto

$$\partial\phi(x) = \{\xi \in \mathbb{R}^n : \phi^0(x;v) \leq \langle \xi, v \rangle, \forall v \in \mathbb{R}^n\}.$$

Para obtermos condições de otimalidade em termos destes conceitos, necessitaremos relembrar as noções de vetor tangente e de vetor normal.

Inicialmente observamos que se C é um subconjunto não vazio de $\mathbb{R}^n$, então a função distância $d_C : \mathbb{R}^n \to \mathbb{R}$ definida por $d_C(x) = \inf_{y \in C} ||y - x||$ não é diferenciável, mas é globalmente Lipschitz - neste caso, com constante de Lipschitz igual a 1.

Sejam $C \subset \mathbb{R}^n$ e $x \in C$. Dizemos que $v \in \mathbb{R}^n$ é um vetor tangente a C em x se $d_C^0(x;v) = 0$. O conjunto dos vetores tangentes será denotado $T_C(x)$ e é chamado **cone tangente** (de Clarke). Por polaridade, se define o **cone normal** (de Clarke),

$$N_C(x) = \{\xi \in \mathbb{R}^n : \langle \xi, v \rangle \leq 0, \forall v \in T_C(x)\}.$$

Pode-se demonstrar que estes cones coincidem com os cones tangente e normal (usuais, da Análise Convexa), quando C é um conjunto convexo. Além disto, para cada $x \in C$, estes conjuntos são, ambos, cones convexos e fechados.

Recordamos ainda um resultado que estabelece uma condição necessária de otimalidade em termos de uma condição estacionária.

Proposição 2.1. *Sejam $\phi : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ uma função definida no conjunto aberto Ω e C um subconjunto não vazio de Ω. Se x^* é um minimizador de ϕ em C, então vale a inclusão*

$$0 \in \partial\phi(x^*) + N_C(x^*). \tag{1}$$

(Além disto, um ponto $x^* \in C$ que satisfaz (1) é chamado ponto estacionário de ϕ em C)

E por fim, lembramos os seguintes resultados.

Lema 2.1. ***(Regra da Cadeia)*** *Sejam $g : \mathbb{R}^n \to \mathbb{R}^k$ uma função estritamente diferenciável em $\hat{y}$ e $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz perto de $\hat{z} = g(\hat{y})$. Então,*

$$\partial(f \circ g)(\hat{y}) \subset \bigcup_{\bar{z} \in \partial f(\hat{z})} \langle \bar{z}, \nabla g(\hat{y}) \rangle.$$

Consideremos o problema de programação matemática multiobjetivo:

$$\begin{array}{l} \text{Minimizar } (f_0(y), \cdots, f_{r_0}(y)) \\ \text{sujeito a:} \\ f_j(y) \leq 0, \; j = r_0 + 1, \cdots, r_0 + r_1 \\ f_j(y) = 0, \; j = r_0 + r_1 + 1, \cdots, r_0 + r_1 + r_2 \\ y \in \Omega \end{array} \tag{2}$$

onde $f_j : \mathbb{R}^n \to \mathbb{R}$, $j = 0, \cdots, r_0 + r_1 + r_2$ e Ω é um subconjunto não vazio de $\mathbb{R}^n$.

As condições necessárias de Fritz John para o problema não diferenciável (2) são dadas no seguinte lema:

DOI: 10.5540/03.2018.006.02.0307

Lema 2.2. *Suponha que $\hat{y}$ é uma solução Pareto fraca do problema (2). Suponha ainda que Ω é um subconjunto convexo, fechado e não vazio de $\mathbb{R}^n$ e que as funções f_j, $j = 0, \cdots, r_0+r_1+r_2$ são Lipschitz perto de $\hat{y}$. Então existem números λ_j, $i = 0, \cdots, r_0+r_1+r_2$, não todos nulos e tais que*

$$\begin{aligned}
&0 \in \sum_{i=0}^{r_0+r_1+r_2} \partial(\lambda_j f_j)(\hat{y}) + N_\Omega(\hat{y}) \\
&\lambda_j \geq 0,\ j = 0, \cdots, r_0 + r_1 \\
&\lambda_j f_j(\hat{y}) = 0,\ j = r_0 + 1, \cdots, r_0 + r_1
\end{aligned}$$

(onde $N_\Omega(\hat{y})$ é o cone normal (usual da Análise Convexa)).

Para mais detalhes sobre Análise Não-diferenciável de Clarke, consulte por exemplo [4].

3 Resultado Principal

Nesta Seção obteremos uma prova simplificada para o Princípio do Máximo Discreto para o Problema de Controle Ótimo Discreto Multiobjetivo Não Diferenciável (PND).

Teorema 3.1. *Suponha que $\varphi_i(\cdot,\cdot)$ é contínua ao redor do processo Pareto fraco $(\hat{x}, \hat{u})$ de (PND); as funções $\varphi_i(\cdot, \hat{u}_i)$ são diferenciáveis em $\hat{x}_i$, para todo $i = 0, \cdots, N-1$; as funções Φ_j, $j = 0, \cdots, r_0 + r_1 + r_2$ são Lipschitz em $(\hat{x}_1, \cdots \hat{x}_N)$ e os conjuntos $\varphi_i(\hat{x}_i, U_i)$ são convexos, para cada $i = 0, \cdots, N-1$. Então, existem*

1. $\lambda_j \in \mathbb{R}$, $j = 0, \cdots, r_0 + r_1 + r_2$, *não todos nulos e tais que*

$$\begin{aligned}
&\lambda_j \geq 0,\ j = 0, \cdots, r_0 + r_1; \\
&\lambda_j \Phi_j(\hat{x}_1, \cdots, \hat{x}_N) = 0,\ j = r_0 + 1, \cdots, r_0 + r_1.
\end{aligned} \tag{3}$$

2. $\bar{z}^j = (\bar{z}^j_1, \cdots, \bar{z}^j_N) \in \mathbb{R}^N$, $\bar{z}^j \in \partial(\lambda_j \Phi_j)(\hat{x}_1, \cdots \hat{x}_N)$, $j = 0, \cdots, r_0 + r_1 + r_2$;

3. $p \in R^{n(N+1)}$ *satisfaz a equação adjunta*

$$\begin{aligned}
&p_i = \nabla_x \mathbb{H}_i(p_{i+1}, \hat{x}_i, \hat{u}_i) - \sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_i,\ i = 0, \cdots, N-1 \\
&p_N = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_N
\end{aligned} \tag{4}$$

(sendo $\mathbb{H}_i(p, x, u) := \langle p_{i+1}, \varphi_i(x_i, u_i) \rangle$). Além disto,

$$\mathbb{H}_i(p_{i+1}, \hat{x}_i, \hat{u}_i) = \max_{u_i \in U_i} \mathbb{H}_i(p_{i+1}, \hat{x}_i, u_i), \textit{para cada } i \in [0, N-1]. \tag{5}$$

A demonstração do Teorema 3.1 é feita mediante a reformulação de (PND) como um problema de programação matemática multiobjetivo, ao qual aplicaremos a Regra de Multiplicadores de Lagrange, Lema 2.2; também utilizaremos a versão da Regra da

DOI: 10.5540/03.2018.006.02.0307

Cadeia, para funções localmente Lipschitz, em termos dos gradientes generalizados de Clarke, Lema 2.1.

Esboço da demonstração **[do Teorema 3.1]** Fixado $\theta \in [0, N-1]$ e $v \in U_\theta$, considere o controle

$$u_i = \begin{cases} \hat{u}_i, \ i \neq \theta \\ v, \ i = \theta. \end{cases}$$

Sejam: $y = \varphi_\theta(\hat{x}_\theta, v)$; $\hat{y} = \varphi_\theta(\hat{x}_\theta, \hat{u}_\theta)$ e x^y a trajetória correspondente ao controle u.

Como $(\hat{x}, \hat{u})$ é um processo ótimo Pareto fraco de (PND), então $y = \hat{y}$ é solução Pareto fraca do seguinte problema multiobjetivo:

$$\begin{array}{l} \text{Minimizar } (\Phi_0(x_1^y, \cdots, x_N^y), \cdots, \Phi_{r_0}(x_1^y, \cdots, x_N^y)) \\ \text{sujeito a:} \\ \Phi_j(x_1^y, \cdots, x_N^y) \leq 0, \ j = r_0 + 1, \cdots, r_0 + r_1 \\ \Phi_j(x_1^y, \cdots, x_N^y) = 0, \ j = r_0 + r_1 + 1, \cdots, r_0 + r_1 + r_2 \\ y \in \varphi_\theta(\hat{x}_\theta, U_\theta). \end{array} \tag{6}$$

Aplicando o Lema 2.2 ao problema (6), obtemos que existem $\lambda_j \in \mathbb{R}$, $j = 0, \cdots, r_0 + r_1 + r_2$, tais que $\lambda_j \geq 0$, $j = 0, \cdots, r_0 + r_1$ e

$$0 \in \sum_{j=0}^{r_0+r_1+r_2} \partial_y(\lambda_j \Phi_j)(x_1^{\hat{y}}, \cdots, x_N^{\hat{y}}) + N_{\varphi_\theta(\hat{x}_\theta, U_\theta)}(\hat{y}) \tag{7}$$

(onde, é claro, $x^{\hat{y}} = \hat{x}$).

A função $y \mapsto x_i^y$ $(i = 0, \cdots, N)$ é diferenciável em $\hat{y}$. Neste caso, a derivada é dada por

$$\nabla_y x_i^{\hat{y}} = \begin{cases} 0, \ i \leq \theta \\ f(\theta + 1, i), \ i \geq \theta + 1 \end{cases}$$

sendo

$$f(s, i) = \begin{cases} \mathbf{I}, \text{ se } s = i \\ \nabla_x \varphi_{i-1}(\hat{x}_{i-1}, \hat{u}_{i-1}) \cdot ... \cdot \nabla_x \varphi_s(\hat{x}_s, \hat{u}_s), \ s < i \end{cases}$$

onde $\mathbf{I}$ é a matriz identidade $n \times n$.

De (7) segue que existem vetores $\bar{\xi} \in N_{\varphi_\theta(\hat{x}_\theta, U_\theta)}(\hat{y})$ e $\bar{y}^j \in \partial_y(\lambda_j \Phi_j)(x_1^{\hat{y}}, \cdots, x_N^{\hat{y}})$, $j = 0, \cdots, r_0 + r_1 + r_2$ e tais que

$$0 = \bar{\xi} + \sum_{j=0}^{r_0+r_1+r_2} \bar{y}^j \tag{8}$$

e, pela definição de cone normal,

$$\langle \bar{\xi}, y - \hat{y} \rangle \leq 0, \forall y \in \varphi_\theta(\hat{x}_\theta, U_\theta) \tag{9}$$

e portanto, de (8) e (9), temos

$$\sum_{j=0}^{r_0+r_1+r_2} \langle \bar{y}^j, y - \hat{y} \rangle \geq 0, \ \forall y \in \varphi_\theta(\hat{x}_\theta, U_\theta). \tag{10}$$

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 6, n. 2, 2018.

Aplicando-se a Regra da Cadeia (Lema 2.1) à função composta $g_j(y) = \Phi_j(x^y)$:

$$\langle \bar{y}^j, y - \hat{y} \rangle = \sum_{s=\theta+1}^{N} \langle \bar{z}_s^j, f(\theta+1, s)(y - \hat{y}) \rangle, \; j = 0, \cdots, r_0 + r_1 + r_2. \tag{11}$$

Defina:

$$p_i := -\sum_{j=0}^{r_0+r_1+r_2} \sum_{s=i}^{N} \bar{z}_s^j f(i, s). \tag{12}$$

Segue de (10) e (11),

$$\langle p_{\theta+1}, y - \hat{y} \rangle = \langle p_{\theta+1}, \varphi_\theta(\overline{x}_\theta, v) - \varphi_\theta(\hat{x}_\theta, \hat{u}_\theta) \rangle \leq 0$$

que é a afirmação do Princípio de Máximo dada por (5). Resta provar que p_i definido em (12) satisfaz a equação adjunta (4). Isto será verificado indutivamente onde admitiremos, por simplicidade, que $\varphi_i(\hat{x}_i, \hat{u}_i) = \varphi_i$.

De fato, tomando-se $i = N$ em (12):

$$p_N = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_N^j \underbrace{f(N, N)}_{=\mathbf{I}} = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_N^j.$$

Agora, faça $i = N - 1$ em (12):

$$p_{N-1} = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-1}^j + \langle p_N, \nabla_x \varphi_{N-1} \rangle = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-1}^j + \nabla_x H_{N-1}(p_N, \hat{x}_{N-1}, \hat{u}_{N-1})$$

e portanto (4) vale para $i = N - 1$.

Tomemos agora $i = N - 2$ em (12):

$$p_{N-2} = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-2}^j + \nabla_x \varphi_{N-2} \left[\bar{z}_{N-1}^j + \bar{z}_N^j \nabla_x \varphi_{N-1} \right]. \tag{13}$$

E por alguns algebrismos temos:

$$-\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-1}^j \nabla_x \varphi_{N-2} - \sum_{j=0}^{r_0+r_1+r_2} \bar{z}_N^j \nabla_x \varphi_{N-1} \nabla_x \varphi_{N-2} = \langle p_{N-1}, \nabla_x \varphi_{N-2} \rangle. \tag{14}$$

Comparando (13) e (14), obtemos

$$p_{N-2} = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-2}^j + \langle p_{N-1} \nabla_x \varphi_{N-2} \rangle = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-2}^j + \nabla_x H_{N-2}(p_{N-1}, \hat{x}_{N-2}, \hat{u}_{N-2}).$$

E, indutivamente, p_i definida em (12) satisfaz a equação adjunta (4), para todo $i = 0, \cdots, N - 1$.

4 Conclusões

Neste trabalho, obtivemos uma versão do Princípio do Máximo para uma certa classe de problemas multiobjetivos de controle ótimo discreto não-diferenciáveis. O resultado foi obtido utilizando resultados de Análise Não-Suave (Gradientes e Derivadas generalizadas de Clarke). O prosseguimento natural deste trabalho é investigar a suficiência destas condições para a otimalidade global e, também, obter resultados similares utilizando alguma noção mais geral de subgradiente.

Referências

[1] A. V. Arutyunov e B. Marinkovic. Necessary conditions of optimality for discrete optimal control problems. *University Computational Mathematics and Cybernetics*, Moscow, 1:38-44, 2005.

[2] R. E. Bellman. *Dynamic Programming*. Mineola, NY: Dover Publications, 2003.

[3] V. G. Boltyanskii. *Optimal Control of Discrete Systems*. Halsted Press, 1978.

[4] F. H. Clarke, Y. S. Ledyaev, R. J. Stern e P. R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer Science & Business Media, volume 178, 2008.

[5] R. Hilscher e V. Zeidan. Second order sufficiency criteria for a discrete optimal control problem. *The Journal of Difference Equations and Applications*, 8-6:573-603, 2002.

[6] B. Marinkovic. Optimality conditions in discrete optimal control problems with state constraints. *Numerical Functional Analysis and Optimization*, volume 28, 7-8:945-955, 2007.

[7] B. Marinkovic. Second-order optimality conditions in a discrete optimal control problem. *Optimization*, 57-4:539-548, 2008.

[8] B. S. Mordukhovich, I. Shvartsman. Discrete maximimum principle for nonsmooth optimal control problems with delays. *Cybernetics and Systems Analysis*, 38:255-264, 2002.

[9] L. Pontryagin, V. Boltianski, R. Gramkrelidze e E. Mitchenko. *Theorie Mathématique des Processus Optimaux,* Mir, Moscow, Russia, 1974.

[10] A. I. Propoi. *Elements of the theory of optimal discrete processes.* M.: Nauka, 1973.

[11] I. Shvartsman. Necessary optimality conditions in discrete nonsmooth optimal control. *Journal of Optimization Theory and Applications*, 153:578-586, 2012.

DOI: 10.5540/03.2018.006.02.0307