



برنامه‌ریزی پویای تقریبی مبتنی بر بهینه‌سازی مجموع مربعات برای سیستم‌های متغیر با زمان و کاربرد آن در طراحی قانون هدایت زیربینه

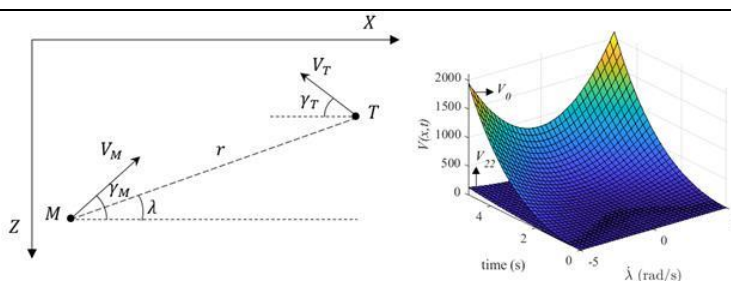
سجاد پاک‌خصال^۱، سعید شمقدری^{۲*}

^۱ دانشجوی دکتری، گروه کنترل، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران
^۲ دانشیار، گروه کنترل، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران

برجسته‌ها

- یک الگوریتم تکرار سیاست برای کنترل زیربینه سیستم‌های متغیر با زمان ارائه شده است.
- خواص مربوط به این الگوریتم اثبات شده‌اند.
- یک قانون هدایت بر اساس این الگوریتم طراحی شده است.

چکیده گرافیکی



مشخصات مقاله

تاریخچه مقاله:
نوع مقاله: علمی پژوهشی
دریافت: ۱۴۰۱/۰۴/۱۹
بازنگری: ۱۴۰۱/۰۵/۱۱
پذیرش: ۱۴۰۱/۰۶/۰۲
ارائه برخط: ۱۴۰۱/۰۷/۲۶
*نویسنده مسئول:
shamaghdari@iust.ac.ir

کلیدواژه‌ها:

قانون هدایت زیربینه
برنامه‌ریزی پویای تقریبی
سیستم متغیر با زمان
بهینه‌سازی مجموع مربعات
الگوریتم تکرار سیاست

چکیده

در این مقاله، روشی برای کنترل زیربینه سیستم‌های چندجمله‌ای متغیر با زمان ارائه و از آن برای طراحی قانون هدایت ره‌گیرها استفاده می‌شود. از آنجایی که معادلات برخورد ره‌گیر و هدف به فاصله بین آن‌ها وابسته هستند و این فاصله در طول پرواز تغییر می‌کند، طراح قانون هدایت با یک سیستم متغیر با زمان مواجه است. روش‌های توسعه داده شده برای کنترل سیستم‌های نامتغیر با زمان، به‌طور مستقیم قابل استفاده برای سیستم‌های متغیر با زمان نیستند. یکی از رویکردهای کنترلی مرسوم برای طراحی قانون هدایت ره‌گیرها، کنترل بهینه می‌باشد. برنامه‌ریزی پویای تقریبی یک روش شناخته شده برای حل مسئله کنترل بهینه است. یکی از چالش‌های کاربرد این روش برای کنترل سیستم‌های غیرخطی متغیر با زمان، سخت بودن حل معادله بلمن است. در روش پیشنهادی این مقاله، حل معادله بلمن با حل یک مسئله بهینه‌سازی مجموع مربعات جایگزین شده است. ثابت می‌شود که سیاست کنترلی طراحی شده با این روش، پایدارساز نمایی فراگیر و زیربینه خواهد بود. درنهایت، کارایی روش پیشنهادی برای هدایت ره‌گیرها، از طریق شبیه‌سازی‌های عددی نشان داده می‌شود.

۱- مقدمه

می‌سازد [۷]. در مورد برنامه‌ریزی پویا، بلمن^۹ این مشکل محاسباتی را نفرین ابعاد^{۱۰} نامیده است [۸]؛ بنابراین از همان سال‌های نخست بسیاری از محققین به‌جای حل بهینه دقیق، روی حل تقریبی مسئله کنترل بهینه تمرکز کرده‌اند [۹ و ۱۰]. برنامه‌ریزی پویای تقریبی^{۱۱} [۱۱] یک روش تکراری برای حل تقریبی مسئله کنترل بهینه می‌باشد و طی چند دهه اخیر توجه بسیاری از محققین را به خود جلب کرده است [۱۲].

دو الگوریتم اصلی در حوزه برنامه‌ریزی پویای تقریبی را می‌توان تکرار سیاست^{۱۲} و تکرار ارزش^{۱۳} دانست. به دلیل مزیت‌هایی که در اثبات پایداری حلقه بسته برای الگوریتم تکرار سیاست وجود دارد، اغلب رویکردهای کنترلی مبتنی بر برنامه‌ریزی پویای تقریبی از این الگوریتم استفاده می‌کنند [۱۲]. دو گام اصلی در الگوریتم‌های برنامه‌ریزی پویای تقریبی، ارزیابی سیاست^{۱۴} و بهبود سیاست^{۱۵} هستند. در طراحی کنترل بهینه برای سیستم‌های غیرخطی با استفاده از الگوریتم تکرار سیاست، در گام ارزیابی سیاست می‌بایست یک معادله دیفرانسیل با مشتقات جزئی^{۱۶} (PDE) حل شود. این معادله به معادله بلمن^{۱۷} معروف است. معادله بلمن نسبت به مجهول خود، یعنی تابع ارزش^{۱۸}، خطی است اما حل PDEهای خطی نیز مشکل است؛ بنابراین معمولاً از روش‌های تقریبی برای به دست آوردن تابع ارزش استفاده می‌شود که معروف‌ترین آن‌ها شبکه‌های عصبی^{۱۹} می‌باشند [۱۲].

پیش‌از این در [۱۳] از ترکیب برنامه‌ریزی پویای تقریبی و مفهوم کنترل H_∞ برای طراحی قانون هدایت استفاده شده است. تقریب تابع ارزش در این مرجع با استفاده از شبکه عصبی انجام شده است. متأسفانه تقریب‌های مبتنی بر شبکه‌های عصبی تنها روی یک زیرمجموعه از فضای حالت

تاکنون از روش‌های مختلفی برای هدایت ره‌گیرها^۱ استفاده شده است که هدایت ناوبری تناسبی^۲ محبوب‌ترین آن‌ها می‌باشد [۱]. ایده اصلی استفاده شده در هدایت ناوبری تناسبی این است که اگر نرخ چرخش خط دید^۳ ره‌گیر و هدف صفر باشد، برخورد بین آن‌ها اتفاق می‌افتد. محبوبیت این روش هدایت از کارایی خوب و کم بودن اطلاعات موردنیاز برای پیاده‌سازی آن نشأت می‌گیرد. این محبوبیت سبب شده است تا محققین بسیاری سعی بر ارتقای این روش هدایت داشته باشند. استفاده از رویکردهای مختلف موجود در نظریه کنترل در ترکیب با ایده ناوبری تناسبی طی سالیان گذشته موضوع تحقیقات متعددی در زمینه هدایت ره‌گیرها بوده است. از جمله این رویکردها می‌توان به کنترل بهینه [۲]، کنترل پیش‌بین مدل^۴ [۳]، کنترل تناسبی-انتگرالی^۵ (PI) [۴]، کنترل مد لغزشی^۶ [۵] و کنترل H_∞ [۶] اشاره کرد.

همانند بسیاری از مسائل مورد مطالعه در نظریه کنترل، در طراحی قانون هدایت نیز طراحان با یک مصالحه بین کارایی و هزینه کنترلی مواجه هستند. از دست دادن کارایی می‌تواند به خطای برخورد منجر شود. از سوی دیگر، سیگنال کنترلی حلقه هدایت، شتاب فرمان داده شده به ره‌گیر می‌باشد. بزرگ بودن این سیگنال نیز در عمل می‌تواند باعث ناتوانی ره‌گیر در ردیابی فرمان و در نتیجه خطای برخورد شود. در کنترل بهینه با تعریف یک تابع هزینه مناسب و کمینه‌سازی آن می‌توان به مصالحه مورد نظر بین کارایی و هزینه کنترلی دست پیدا کرد.

دو رویکرد اصلی حل مسئله کنترل بهینه، رویکرد مبتنی بر حساب تغییرات^۷ و برنامه‌ریزی پویا^۸ هستند؛ اما هر دو این روش‌ها از لحاظ محاسباتی با مشکلاتی جدی روبرو هستند که کاربرد آن‌ها را برای حل یک مسئله عمومی ناممکن

9 Bellman

10 Curse of dimensionality

11 Approximate dynamic programming

12 Policy iteration

13 Value iteration

14 Policy evaluation

15 Policy improvement

16 Partial differential equation

17 Bellman equation

18 Value function

19 Neural networks

1 Pursuits guidance

2 Proportional navigation guidance

3 Line-of-sight

4 Model predictive control

5 Proportional-integral control

6 Sliding mode control

7 Calculus of variations

8 Dynamic programming

چندجمله‌ای^۷ متغیر با زمان به روش برنامه‌ریزی پویای تقریبی ارائه شده است. در این مرجع حل معادله بلمن با حل یک مسئله برنامه‌ریزی مجموع مربعات^۸ (SOS) [۱۵] جایگزین شده است. برنامه‌ریزی مجموع مربعات ابزاری برای حل مسائل بهینه‌سازی با قیود ناتساوی چندجمله‌ای و تابع هدف خطی می‌باشد. حل مسئله بهینه‌سازی مجموع مربعات با حل یک مسئله برنامه‌ریزی نیمه‌معین^۹ معادل است [۱۵]؛ بنابراین روش ارائه شده در [۱۴] مشکلی از لحاظ حجم محاسبات نخواهد داشت. با توجه به اینکه هنگام حل مسئله بهینه‌سازی بازه محدودکننده‌ای از فضای حالت در نظر گرفته نمی‌شود، اثبات‌های انجام شده در [۱۴] محلی نمی‌باشند.

از دیگر مقالاتی که به مطالعه کنترل زیربهینه سیستم‌های چندجمله‌ای با رویکرد برنامه‌ریزی مجموع مربعات پرداخته‌اند می‌توان به مراجع [۱۶-۱۸] اشاره کرد. در [۱۶] و [۱۷] به ترتیب الگوریتم‌هایی باهدف طراحی کنترل زیربهینه برای سیستم‌های نامتغیر با زمان چندجمله‌ای زمان پیوسته و زمان گسسته ارائه شده‌اند. در [۱۸] نیز الگوریتمی باهدف تضمین عدم ورود حالت‌های سیستم حلقه بسته به ناحیه‌ای از پیش تعیین شده ارائه شده است. این روش نیز مخصوص سیستم‌های نامتغیر با زمان چندجمله‌ای زمان پیوسته بوده و اصطلاحاً کنترل زیربهینه ایمن نامیده شده است.

ساختار یک مسئله بهینه‌سازی این امکان را به طراح می‌دهد تا به منظور بهبود کارایی، ایده‌های مختلفی را با کمک تغییر در قیود، تغییر در تابع هدف و افزودن قیود جدید، به الگوریتم اعمال کند. در این مقاله، مسئله بهینه‌سازی ارائه شده در [۱۴] باهدف تضمین پایداری نمایی^{۱۰} سیستم حلقه بسته تغییر داده می‌شود. سپس از الگوریتم پیشنهاد شده برای طراحی قانون هدایت بهینه در فاز آشیانه‌یاب ره‌گیرها استفاده می‌شود. به این صورت که با فرض ثابت بودن سرعت نزدیک‌شوندگی^{۱۱} در فاز آشیانه‌یاب،

معتبر هستند. این خاصیت باعث می‌شود که تحلیل‌های انجام شده تنها به صورت محلی^۱ معتبر باشند. همچنین در صورت نیاز به افزایش دقت تقریب می‌بایست از تعداد زیادی توابع پایه^۲ استفاده شود. این کار می‌تواند منجر به مشکلات محاسباتی شود. علاوه بر این، در تقریب‌های مبتنی بر شبکه‌های عصبی نمی‌توان مثبت بودن تابع منتجه را تضمین کرد. این در حالی است که مثبت بودن برای تابع ارزش امری ضروری است، زیرا در هنگام اثبات پایداری از آن به عنوان تابع لیاپانوف^۳ استفاده می‌شود. نکته قابل توجه در طراحی قانون هدایت این است که معادله دیفرانسیلی نرخ چرخش خط دید به فاصله ره‌گیر تا هدف وابسته است و این فاصله در طول پرواز در حال تغییر است؛ بنابراین، طراح هدایت با یک سیستم متغیر با زمان مواجه خواهد بود. در الگوریتم ارائه شده در [۱۳]، همانند تعداد زیادی از الگوریتم‌های طراحی هدایت مبتنی بر کنترل مدرن، از متغیر با زمان بودن فاصله ره‌گیر و هدف صرف نظر شده است. این مرجع، در طول فاز نهایی^۴ (آشیانه‌یاب^۵) پرواز از یک مقدار ثابت کوچک به عنوان فاصله ره‌گیر تا هدف استفاده می‌کند که این کار سبب افزایش محافظه‌کاری در طراحی انجام شده می‌شود.

باید به این نکته توجه داشت که استراتژی‌های توسعه داده شده برای کنترل سیستم‌های نامتغیر با زمان به سادگی قابل استفاده برای سیستم‌های متغیر با زمان نیستند. تفاوت بین کنترل بهینه برای سیستم‌های متغیر با زمان و نامتغیر با زمان به تفاوت شکل معادله‌های بلمن برای این دو نوع سیستم برمی‌گردد. در سیستم‌های متغیر با زمان، سیاست کنترلی و تابع ارزش هر دو وابسته به زمان هستند. این وابستگی به زمان، باعث پیچیدگی بیشتر معادله بلمن و در نتیجه سخت شدن دستیابی به تقریب تابع ارزش می‌شود. اخیراً در [۱۴] روشی توسط نویسندگان مقاله حاضر به منظور طراحی کنترل زیربهینه^۶ برای سیستم‌های

¹ Local

² Basis function

³ Lyapunov

⁴ Terminal phase

⁵ Homing

⁶ Sub-optimal

⁷ Polynomial

⁸ Sum-of-squares programming

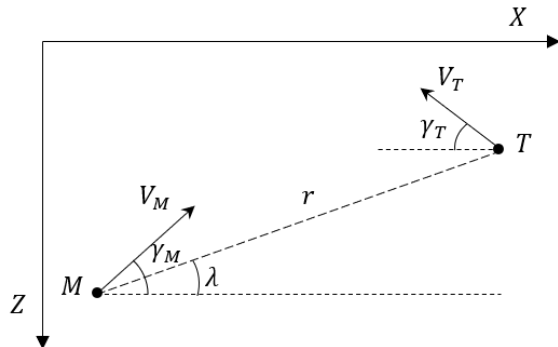
⁹ Semidefinite programming

¹⁰ Exponential stability

¹¹ Closing velocity

۲-۱- سینماتیک برخورد ره‌گیر و هدف

شکل ۱ نشان‌دهنده هندسه درگیری صفحه‌ای بین یک ره‌گیر و هدف در دستگاه مرجع اینرسی^۱ می‌باشد. خط واصل ره‌گیر و هدف را اصطلاحاً خط دید می‌گویند. در این شکل، ره‌گیر و هدف به ترتیب با M و T نشان داده شده‌اند. r فاصله نسبی بین ره‌گیر و هدف است. V_M و V_T به ترتیب سرعت‌های ره‌گیر و هدف هستند. λ زاویه خط دید و γ_M و γ_T به ترتیب زاویه‌های مسیر پرواز^۲ ره‌گیر و هدف می‌باشند.



شکل (۱): هندسه درگیری صفحه‌ای ره‌گیر و هدف.

سینماتیک درگیری ره‌گیر و هدف می‌تواند با معادلات دیفرانسیل زیر بیان شود:

$$\dot{r} = -V_M \cos(\gamma_M - \lambda) - V_T \cos(\gamma_T - \lambda), \quad (1)$$

$$r\dot{\lambda} = -V_M \sin(\gamma_M - \lambda) + V_T \sin(\gamma_T - \lambda). \quad (2)$$

در معادله (۱)، به \dot{r} سرعت نزدیک‌شوندگی نیز می‌گویند. با مشتق‌گیری از معادله (۲) نسبت به زمان می‌توان نوشت:

$$\ddot{\lambda} = \frac{1}{r}(-2\dot{r}\dot{\lambda} - a_M + a_T), \quad (3)$$

که در آن $a_M = V_M \dot{\gamma}_M \cos(\gamma_M - \lambda)$ و $a_T = V_T \dot{\gamma}_T \cos(\gamma_T - \lambda)$ به ترتیب مؤلفه‌های عمود بر خط دید شتاب ره‌گیر و هدف می‌باشند.

با فرض سرعت نزدیک‌شوندگی ثابت $\dot{r} = -V_c$ در فاز نهایی پرواز می‌توان نوشت:

$$\ddot{\lambda} = \frac{1}{r}(2V_c \dot{\lambda} - a_M + a_T) = \frac{1}{r_0 - V_c t}(2V_c \dot{\lambda} - a_M + a_T), \quad (4)$$

فاصله ره‌گیر تا هدف به صورت توابع چندجمله‌ای بر حسب زمان تقریب زده می‌شود و سپس برای سیستم چندجمله‌ای متغیر با زمان به دست آمده قانون هدایت زیر بهینه طراحی می‌شود.

از جمله نوآوری‌ها و تمایزات مقاله حاضر نسبت به مقالات پیشین و به طور خاص مرجع [۱۴] می‌توان به موارد زیر اشاره کرد:

- سیاست کنترلی حاصل از الگوریتم ارائه شده در این مقاله پایدارساز نمایی می‌باشد.
- با توجه به ارائه یک الگوریتم جدید، خواص آن شامل همگرایی الگوریتم، شدنی بودن مسئله بهینه‌سازی و پایداری نمایی فراگیر سیستم حلقه بسته در قالب یک قضیه بیان و اثبات شده‌اند.
- از الگوریتم ارائه شده، برای طراحی قانون هدایت بهینه که یک مسئله کاربردی می‌باشد، استفاده شده است.

ساختار این مقاله بدین صورت است که پس از مقدمه ارائه شده در این بخش، در بخش دوم پیش‌نیازهای لازم برای درک ادامه مباحث مقاله آورده شده است. بخش دوم شامل سه زیر بخش است که در آن‌ها به ترتیب معادلات برخورد ره‌گیر-هدف، تعاریف و مفاهیم برنامه‌ریزی مجموع مربعات و برنامه‌ریزی پویای تقریبی تشریح می‌شوند. در بخش سوم، الگوریتم تکرار سیاست پیشنهادی معرفی شده و خواص آن در قالب یک قضیه بیان می‌شوند. در بخش چهارم، از الگوریتم پیشنهادی برای طراحی قانون هدایت استفاده شده و با بهره‌گیری از شبیه‌سازی‌های عددی عملکرد این قانون هدایت به اثبات رسیده است. بخش پنجم نیز به بیان نتیجه‌گیری مقاله اختصاص یافته است.

۲- مبانی طراحی قانون هدایت بهینه مبتنی

بر برنامه‌ریزی مجموع مربعات

این بخش، حاوی مبانی موردنیاز برای مطالب ارائه شده در بخش‌های بعدی می‌باشد. این مبانی شامل مدل‌سازی سینماتیکی برخورد ره‌گیر و هدف و تعاریف مربوط به برنامه‌ریزی مجموع مربعات و کنترل بهینه می‌باشند.

¹ Inertial reference frame

² Flight path angle

فرم کلی مسائل بهینه‌سازی مجموع مربعات به صورت زیر است:

$$\min \omega^T c \quad (5)$$

$$\text{s. t. } a_{0,j}(x) + \sum_{i=1}^N p_i(x) a_{i,j}(x) = 0 \quad \text{for } j = 1, \dots, \hat{j} \quad (6)$$

$$a_{0,j}(x) + \sum_{i=1}^N p_i(x) a_{i,j}(x) \text{ is SOS for } j = (\hat{j} + 1), (\hat{j} + 2), \dots, J \quad (7)$$

که c برداری از ضرایب مجهول چندجمله‌ای‌های $p_i(x)$ و $i = 1, 2, \dots, \hat{N}$ و چندجمله‌ای‌های مجموع مربعات $a_{i,j}(x)$ و $i = (\hat{N} + 1), \dots, N$ بردار وزن‌دهی و همچنین x متغیر SOS نامیده می‌شود. SOSTOOLS [۲۰] یک جعبه‌ابزار رایگان در متلب^۳ برای تبدیل مسئله بهینه‌سازی مجموع مربعات به مسئله برنامه‌ریزی نیمه‌معین است. این جعبه‌ابزار امکان استفاده از حل‌کننده‌های مختلفی از جمله Mosek [۲۱] را برای حل مسائل برنامه‌ریزی نیمه‌معین فراهم کرده است.

۳-۲- کنترل بهینه و الگوریتم تکرار سیاست

سیستم متغیر با زمان چندجمله‌ای زیر را در نظر بگیرید:

$$\dot{x}(t) = f(x(t), t) + g(x(t), t)u(x(t), t), \quad (8)$$

که در آن $x(t) \in \mathbb{R}^n$ بردار حالت، $u(x(t), t) \in \mathbb{R}^m$ ورودی کنترلی و f و g نگاشت‌های چندجمله‌ای با فرض $f(0, t) = 0$ هستند. به منظور سادگی نگارش از این پس $x(t)$ به صورت x نوشته می‌شود.

در مسئله کنترل بهینه، هدف یافتن سیاست کنترلی $u(x, t)$ است که تابع ارزش تعریف شده به صورت زیر را کمینه می‌کند:

$$V(x, t) = \int_t^\infty r(x, u(x, \tau), \tau) d\tau. \quad (9)$$

در رابطه فوق $r(x, u(x, t), t) = q(x, t) + Ru(x, t)u^T(x, t)Ru(x, t)$ تابع هزینه لحظه‌ای است که در آن

که در آن r_0 فاصله ره‌گیر تا هدف در شروع فاز نهایی پرواز است.

با بهره‌گیری از ایده ناوبری تناسبی می‌توان از a_M برای کنترل $\hat{\lambda}$ و صفر کردن آن به منظور برخورد ره‌گیر به هدف استفاده کرد. باید به این نکته اشاره کرد که با توجه به روابط اگر شتاب به صورت عمود بر خط دید به ره‌گیر وارد شود اثر بیشتری در صفر کردن $\hat{\lambda}$ خواهد داشت. بنابراین در این مقاله برای کنترل $\hat{\lambda}$ از شتاب عمود بر خط دید به جای شتاب عمود بر سرعت استفاده شده است. به این ایده ناوبری تناسبی حقیقی^۱ نیز گفته می‌شود [۱].

۲-۲- برنامه‌ریزی مجموع مربعات

در این زیر بخش، مفاهیم و تعاریف پایه برنامه‌ریزی مجموع مربعات به طور خلاصه بیان می‌شوند.

حاصل ضرب چند متغیر با نماهای از جنس عدد صحیح مثبت $(m(x) = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n})$ یک تک‌جمله‌ای^۲ نامیده می‌شود. درجه یک تک‌جمله‌ای نیز به صورت مجموع نماهای متغیرهای موجود در آن $(\sum_{i=1}^n \alpha_i)$ تعریف می‌شود. یک چندجمله‌ای نیز ترکیب خطی تعدادی تک‌جمله‌ای با ضرایب حقیقی است. اگر چندجمله‌ای‌های $p_1(x), \dots, p_m(x)$ وجود داشته باشند بطوریکه $p(x) = \sum_{i=1}^m p_i^2(x)$ چندجمله‌ای مجموع مربعات است.

مسائل مختلفی در علوم مهندسی وجود دارند که برای حل آن‌ها نیاز است نامنفی بودن تعدادی چندجمله‌ای بررسی شود؛ اما بررسی نامنفی بودن چندجمله‌ای‌ها با درجه بیشتر از چهار بسیار مشکل است [۱۵]. از طرفی ثابت شده است که بررسی مجموع مربعات بودن یک چندجمله‌ای، معادل با حل یک مسئله برنامه‌ریزی نیمه‌معین است [۱۵]. مسائل برنامه‌ریزی نیمه‌معین نیز می‌توانند به کمک حل‌کننده‌های مربوطه بدون سختی محاسباتی حل شوند [۱۹]. واضح است که هر چندجمله‌ای که بتواند به صورت مجموع مربعات نوشته شود نامنفی است. از این رو در فرمول‌بندی مسائل می‌توان با محافظه‌کاری کمی شرط نامنفی بودن را با شرط کافی مجموع مربعات بودن جایگزین کرد.

¹ True proportional navigation

² Monomial

³ MATLAB

$$u_{i+1}(x, t) = \arg \min_{u(\cdot)} \left\{ r(x, u(x, t), t) + \left(\frac{\partial V_i(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u(x, t)) \right\} = -\frac{1}{2} R^{-1} g^T(x, t) \frac{\partial V_i(x, t)}{\partial x}. \quad (15)$$

البته (۱۴) همچنان یک PDE بوده که به راحتی قابل حل نیست و برای حل آن از روش‌های مختلف تقریبی استفاده می‌شود. متداول‌ترین روش برای تقریب $V_i(x, t)$ استفاده از شبکه عصبی است. البته استفاده از این روش محدودیت‌هایی دارد که در مقدمه این مقاله به آن‌ها اشاره شد. اخیراً توسط نویسندگان مقاله حاضر الگوریتم تکرار سیاستی در [۱۴] ارائه شده است که محدودیت‌های مربوط به استفاده از شبکه عصبی را ندارد. در این مرجع، PDE موجود در گام ارزیابی سیاست با یک مسئله بهینه‌سازی مجموع مربعات جایگزین شده که حل آن از نظر محاسباتی ساده است.

۳- روش کنترلی پیشنهادی

در این بخش، نسخه بهبودیافته‌ای از الگوریتم پیشنهاد شده در [۱۴] ارائه می‌شود. در این الگوریتم باهدف تضمین پایداری نمایی، قیودی به مسئله بهینه‌سازی موجود در گام ارزیابی سیاست اضافه شده است. پیش از ارائه الگوریتم، فرض زیر را در نظر بگیرید:

فرض ۱: $(V_0(x, t), P_0(x, t), u_1(x, t))$ وجود دارد به طوری که

$$-r(x, u_1(x, t), t) - \left(\frac{\partial V_0(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_1(x, t)) - \frac{\partial V_0(x, t)}{\partial t} \text{ is SOS}, \quad (16)$$

$$V_0(x, t) - \varepsilon_1 x^T x \text{ is SOS}, \quad (17)$$

$$\varepsilon_2 P_0(x, t) - V_0(x, t) \text{ is SOS}, \quad (18)$$

$$q(x, t) - \alpha P_0(x, t) \text{ is SOS}, \quad (19)$$

$$P_0(x, t) \text{ is SOS}. \quad (20)$$

الگوریتم تکرار سیاست پیشنهادی به صورت زیر است:

الگوریتم ۲:

۱. مقداردهی اولیه: در نظر بگیرید که $(V_0(x, t), P_0(x, t), u_1(x, t))$ فرض ۱ را ارضا می‌کند.

$q(x, t)$ یک تابع چندجمله‌ای مثبت و R یک ماتریس مثبت معین است.

با در نظر گرفتن سیستم (۸) و مشتق‌گیری از طرفین معادله (۹) می‌توان نوشت:

$$r(x, u(x, t), t) + \left(\frac{\partial V(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u(x, t)) + \frac{\partial V(x, t)}{\partial t} = 0, \quad V(0, t) = 0. \quad (10)$$

تابع ارزش بهینه به صورت زیر تعریف می‌شود:

$$V^0(x, t) = \min_{u(\cdot)} \left\{ \int_t^\infty r(x, u(x, \tau), \tau) d\tau \right\}. \quad (11)$$

بر این اساس سیاست کنترل بهینه را نیز می‌توان به صورت زیر تعریف کرد:

$$u^0(x, t) = \arg \min_{u(\cdot)} \left\{ \int_t^\infty r(x, u(x, \tau), \tau) d\tau \right\}. \quad (12)$$

با جایگذاری (۱۲) در (۱۰)، معادله دیفرانسیل با مشتقات جزئی زیر به دست می‌آید:

$$-\frac{\partial V^0(x, t)}{\partial t} = r(x, u^0(x, t), t) + \left(\frac{\partial V^0(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u^0(x, t)), \quad V^0(0, t) = 0. \quad (13)$$

معادله فوق به معادله همیلتون-ژاکوبی-بلمن^۱ زمان پیوسته معروف بوده و حل تحلیلی آن در حالت عمومی غیرممکن است. معروف‌ترین روش تقریبی برای حل این معادله استفاده از الگوریتم تکرار سیاست به صورت زیر است:

الگوریتم ۱:

۱. مقداردهی اولیه: فرض کنید $u_1(x, t)$ یک کنترل‌کننده پایدارساز فراگیر^۲ باشد. برای $i = 1, 2, \dots$ مراحل زیر را تا رسیدن به همگرایی انجام دهید:

۲. ارزیابی سیاست: با در نظر گرفتن سیاست کنترل $u_i(x, t)$ معادله بلمن زیر را برای به دست آوردن $V_i(x, t)$ حل کنید:

$$r(x, u_i(x, t), t) + \left(\frac{\partial V_i(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_i(x, t)) + \frac{\partial V_i(x, t)}{\partial t} = 0, \quad V(0, t) = 0. \quad (14)$$

۳. بهبود سیاست: سیاست کنترل را با استفاده از رابطه زیر به‌روزرسانی کنید:

¹ Hamilton-Jacobi-Bellman

² Global

۴. دنباله $\{V_i(x, t)\}$ همگراست و ناساوی زیر صادق است:

$$0 \leq V_\infty(x, t) - V^0(x, t) \leq - \int_t^\infty \left[\left(\frac{\partial V_\infty(x, \tau)}{\partial x} \right)^T (f(x, \tau) + g(x, \tau)u_\infty(x, \tau)) + \frac{\partial V_\infty(x, \tau)}{\partial \tau} + r(x, u_\infty, \tau) \right] d\tau, \quad (29)$$

که در آن $V_\infty(x, t) = \lim_{i \rightarrow \infty} V_i(x, t)$ و $u_\infty(x, t) = -\frac{1}{2}R^{-1}g^T(x, t)\frac{\partial V_\infty(x, t)}{\partial x}$ است.

اثبات:

۱. این گزاره به کمک استقرای ریاضی اثبات می‌شود. طبق فرض ۱، به ازای $i = 1$ یک پاسخ شدنی برای مسئله (۲۷)-(۲۱) است. بنابراین این گزاره به ازای $i = 1$ صحیح است. فرض کنید گزاره به ازای $i = j \geq 2$ صحیح و $(V_j(x, t), P_j(x, t))$ یک حل برای مسئله بهینه‌سازی (۲۷)-(۲۱) به ازای $i = j$ است. حال باید نشان داد که $(V_j(x, t), P_j(x, t))$ یک حل شدنی برای مسئله بهینه‌سازی (۲۷)-(۲۱) به ازای $i = j + 1$ است. با جایگذاری $u_{j+1}(x, t) = -\frac{1}{2}R^{-1}g^T(x, t)\frac{\partial V_j(x, t)}{\partial x}$ و $V_j(x, t)$ در (۲۲) می‌توان نوشت:

$$\begin{aligned} & -u_{j+1}^T(x, t)Ru_{j+1}(x, t) - q(x, t) - \\ & \left(\frac{\partial V_j(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_{j+1}(x, t)) - \\ & \frac{\partial V_j(x, t)}{\partial t} = -u_j^T(x, t)Ru_j(x, t) - q(x, t) - \\ & \left(\frac{\partial V_j(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_j(x, t)) \quad (30) \\ & - \frac{\partial V_j(x, t)}{\partial t} + \\ & (u_j(x, t) - u_{j+1}(x, t))^T R (u_j(x, t) - \\ & u_{j+1}(x, t)). \end{aligned}$$

طبق فرض استقرا رابطه زیر برقرار است:

$$\begin{aligned} & -u_j^T(x, t)Ru_j(x, t) - q(x, t) - \\ & \left(\frac{\partial V_j(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_j(x, t)) - \\ & \frac{\partial V_j(x, t)}{\partial t} \text{ is SOS.} \quad (31) \end{aligned}$$

از این رو:

$$\begin{aligned} & -u_{j+1}^T(x, t)Ru_{j+1}(x, t) - q(x, t) - \\ & \left(\frac{\partial V_j(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_{j+1}(x, t)) - \\ & \frac{\partial V_j(x, t)}{\partial t} \text{ is SOS.} \quad (32) \end{aligned}$$

برای $i = 1, 2, \dots$ مراحل زیر را تا رسیدن به همگرایی انجام دهید:

۲. ارزیابی سیاست: با در نظر گرفتن سیاست کنترل $u_i(x, t)$ مسئله زیر را برای به دست آوردن $V_i(x, t)$ و $P_i(x, t)$ حل کنید:

$$\min \int_{x \times T} V_i(x, t) dx dt \quad (21)$$

$$\begin{aligned} \text{s. t. } & -r(x, u_i(x, t), t) - \\ & \left(\frac{\partial V_i(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_i(x, t)) - \\ & \frac{\partial V_i(x, t)}{\partial t} \text{ is SOS} \end{aligned} \quad (22)$$

$$V_{i-1}(x, t) - V_i(x, t) \text{ is SOS} \quad (23)$$

$$V_i(x, t) - \varepsilon_1 x^T x \text{ is SOS} \quad (24)$$

$$\varepsilon_2 P_i(x, t) - V_i(x, t) \text{ is SOS} \quad (25)$$

$$q(x, t) - \alpha P_i(x, t) \text{ is SOS} \quad (26)$$

$$P_i(x, t) \text{ is SOS.} \quad (27)$$

۳. بهبود سیاست: سیاست کنترل را با استفاده از رابطه زیر به روزرسانی کنید:

$$\begin{aligned} u_{i+1}(x, t) = \arg \min_{u(\cdot)} \{ & r(x, u(x, t), t) + \\ & \left(\frac{\partial V_i(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u(x, t)) \} = \\ & -\frac{1}{2}R^{-1}g^T(x, t)\frac{\partial V_i(x, t)}{\partial x}. \end{aligned} \quad (28)$$

در رابطه (۲۱)، $\mathcal{X} \subset \mathbb{R}^n$ و $\mathcal{T} \subset \mathbb{R}$ مجموعه‌هایی بسته و دلخواه هستند. در مسئله استاندارد بهینه‌سازی مجموع مربعات، که در زیر بخش ۲-۲ معرفی شد، تابع هدف به صورت خطی می‌باشد. انتگرال‌گیری انجام‌شده در رابطه (۲۱) کمک می‌کند که تابع هدف مسئله بهینه‌سازی ارائه‌شده، به صورت خطی نسبت به متغیرهای تصمیم تبدیل شود.

چند خاصیت مربوط به الگوریتم ۲ در قضیه زیر آورده شده‌اند.

قضیه ۱: موارد زیر در مورد الگوریتم ۲ صادق هستند:

۱. مسئله بهینه‌سازی (۲۷)-(۲۱) یک مجموعه شدنی ناتهی دارد.

۲. سیستم حلقه بسته متشکل از (۸) و $u(x, t) = u_{i+1}(x, t)$ در مبدأ پایدار نمایی فراگیر است.

۳. $V_i(x, t) \geq V_{i+1}(x, t) \geq 0$.

نهایتاً با توجه به قید (۲۵) ناتساوی فوق را می‌توان به صورت زیر نوشت:

$$\|x\|_2 \leq \frac{\sqrt{\varepsilon_2 P(x(0))}}{\varepsilon_1} e^{-\frac{\alpha t}{2\varepsilon_2}}. \quad (43)$$

طبق تعریف ارائه‌شده در [۲۲]، اگر ثابت مثبت c ، k و λ وجود داشته باشند به طوری که

$$\|x(t)\| \leq k\|x(0)\|e^{-\lambda t} \quad (43)$$

آنگاه سیستم $\dot{x} = f(x, t)$ در مبدأ پایدار نمایی فراگیر است. بر اساس این تعریف و ناتساوی (۴۳)، سیستم حلقه بسته در مبدأ پایدار نمایی فراگیر است.

۳. با توجه به قیود (۲۳) و (۲۴) این گزاره بدیهی است.

۴. طبق بند ۳، $\{V_i(x, t)\}$ یک دنباله کاهشی است که حد پایینی برابر با صفر دارد. بنابراین دنباله $\{V_i(x, t)\}$ همگرا بوده و یک تابع ارزش حدی به صورت $V_\infty(x, t) = \lim_{i \rightarrow \infty} V_i(x, t)$ وجود دارد. از آنجایی که $V_\infty(x, t)$ یک حل شدنی برای مسئله بهینه‌سازی (۲۷)–(۲۱) است، با توجه به قید (۲۲) ناتساوی زیر برقرار است:

$$\begin{aligned} & -r(x, u_\infty(x, t), t) \\ & - \left(\frac{\partial V_\infty(x, t)}{\partial x} \right)^T (f(x, t) \\ & + g(x, t)u_\infty(x, t)) \\ & - \frac{\partial V_\infty(x, t)}{\partial t} \text{ is SOS.} \end{aligned} \quad (45)$$

مشتق تابع ارزش $V_\infty(x, t)$ نسبت زمان به صورت زیر است:

$$\frac{dV_\infty(x, t)}{dt} = \left(\frac{\partial V_\infty(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_\infty(x, t)) + \frac{\partial V_\infty(x, t)}{\partial t}. \quad (46)$$

با انتگرال‌گیری از طرفین رابطه فوق روی بازه $[t, T]$ می‌توان نوشت:

$$\begin{aligned} V_\infty(x(t), t) &= - \int_t^T \left[\left(\frac{\partial V_\infty(x, \tau)}{\partial x} \right)^T (f(x, \tau) + g(x, \tau)u_\infty(x, \tau)) + \frac{\partial V_\infty(x, \tau)}{\partial \tau} \right] d\tau + \\ V_\infty(x(T), T) &= \int_t^T \left[r(x, u_\infty(x, \tau), \tau) - \right. \\ & \left. r(x, u_\infty(x, \tau), \tau) - \left(\frac{\partial V_\infty(x, \tau)}{\partial x} \right)^T (f(x, \tau) + g(x, \tau)u_\infty(x, \tau)) - \frac{\partial V_\infty(x, \tau)}{\partial \tau} \right] d\tau + \\ & V_\infty(x(T), T). \end{aligned} \quad (47)$$

با در نظر داشتن (۴۵) می‌توان ناتساوی زیر را از رابطه (۴۷) نتیجه گرفت:

پس $(V_j(x, t), P_j(x, t))$ یک جواب شدنی برای مسئله بهینه‌سازی (۲۷)–(۲۱) به ازای $i = j + 1$ است. بنابراین طبق اصل استقرای ریاضی گزاره بیان‌شده در بند ۱ به ازای هر عدد طبیعی i ، صادق است.

۲. با استفاده از ناتساوی (۳۲) می‌توان نوشت:

$$\begin{aligned} & \left(\frac{\partial V_i(x, t)}{\partial x} \right)^T (f(x, t) + g(x, t)u_{i+1}(x, t)) \\ & + \frac{\partial V_i(x, t)}{\partial t} \end{aligned} \quad (33)$$

$$\leq -u_{i+1}^T(x, t)Ru_{i+1}(x, t) - q(x, t)$$

$$\rightarrow \frac{dV_i(x, t)}{dt} \leq - (u_{i+1}^T(x, t)Ru_{i+1}(x, t) + q(x, t)). \quad (34)$$

همچنین از رابطه (۲۶) برداشت می‌شود که:

$$\begin{aligned} \alpha P_i(x, t) \leq q(x, t) &\rightarrow \alpha P_i(x, t) \leq \\ u_{i+1}^T(x, t)Ru_{i+1}(x, t) &+ q(x, t). \end{aligned} \quad (35)$$

پس می‌توان نتیجه گرفت:

$$\frac{dV_i(x, t)}{dt} \leq -\alpha P_i(x, t). \quad (36)$$

از طرفی با توجه به قید (۲۵)، می‌توان دریافت که:

$$\frac{V_i(x, t)}{\varepsilon_2} \leq P_i(x, t). \quad (37)$$

بنابراین:

$$\frac{\dot{V}(x, t)}{V_i(x, t)} \leq -\frac{\alpha}{\varepsilon_2}. \quad (38)$$

رابطه (۳۸) را می‌توان به صورت زیر نیز بازنویسی کرد:

$$\frac{d}{dt} (\ln(V_i(x, t))) \leq -\frac{\alpha}{\varepsilon_2}. \quad (39)$$

با انتگرال‌گیری از طرفین ناتساوی فوق روی بازه $[0, t]$ ، می‌توان نوشت:

$$\begin{aligned} \ln(V_i(x, t)) &\leq \ln(V_i(x(0), 0)) - \frac{\alpha}{\varepsilon_2} t \rightarrow \\ V_i(x, t) &\leq V_i(x(0), 0) e^{-\frac{\alpha t}{\varepsilon_2}}. \end{aligned} \quad (40)$$

پس از بازنویسی قید (۲۴) به صورت:

$$\|x\|_2 \leq \frac{\sqrt{V_i(x, t)}}{\varepsilon_1} \quad (41)$$

می‌توان رابطه زیر را از این ناتساوی و رابطه (۴۰) استنباط کرد:

$$\|x\|_2 \leq \frac{\sqrt{V_i(x(0), 0) e^{-\frac{\alpha t}{\varepsilon_2}}}}{\varepsilon_1}. \quad (42)$$

$$\begin{aligned} & \frac{\partial V^O(x,t)}{\partial t} \\ &= \left(\frac{\partial V_{\infty}(x,t)}{\partial x} - \frac{\partial V^O(x,t)}{\partial x} \right)^T (f(x,t) + \\ & g(x,t)u_{\infty}(x,t)) + \frac{\partial V_{\infty}(x,t)}{\partial \tau} - \frac{\partial V^O(x,t)}{\partial \tau} + \\ & (u_{\infty}(x,t) - u^O(x,t))^T R(u_{\infty}(x,t) - \\ & u^O(x,t)) \\ & \geq \left(\frac{\partial V_{\infty}(x,t)}{\partial x} - \frac{\partial V^O(x,t)}{\partial x} \right)^T (f(x,t) + \\ & g(x,t)u_{\infty}(x,t)) + \frac{\partial V_{\infty}(x,t)}{\partial t} - \frac{\partial V^O(x,t)}{\partial t}. \end{aligned}$$

با انتگرال‌گیری از طرفین رابطه (۵۲) ناتساوی زیر به دست می‌آید:

$$\begin{aligned} & V_{\infty}(x,t) - V^O(x,t) \\ & \leq - \int_t^{\infty} \left[\left(\frac{\partial V_{\infty}(x,\tau)}{\partial x} \right)^T (f(x,\tau) \right. \right. \\ & \left. \left. + g(x,\tau)u_{\infty}(x,\tau) + \frac{\partial V_{\infty}(x,\tau)}{\partial \tau} \right. \right. \\ & \left. \left. + r(x,u_{\infty}(x,t),\tau) \right] d\tau \quad \forall t \geq 0. \end{aligned} \quad (53)$$

پس بخش دوم ناتساوی (۲۹) نیز اثبات شد. ■

از آنجایی طبق بند ۴ قضیه ۱، $V_{\infty}(x,t) \geq V^O(x,t)$ است، روش ارائه‌شده طراحی کنترل زیربهبینه محسوب می‌شود.

۴- طراحی قانون هدایت و شبیه‌سازی‌های عددی

در این بخش، با استفاده از الگوریتم پیشنهادشده در بخش ۳ و مدل سینماتیکی معرفی‌شده در زیر بخش ۲-۱، یک قانون هدایت زیربهبینه برای فاز نهایی ره‌گیرها طراحی‌شده است. همچنین به‌منظور بررسی کارایی قانون هدایت طراحی‌شده، شبیه‌سازی‌هایی در قالب دو سناریوی درگیری ارائه می‌شوند.

در هر دو سناریویی که در ادامه مطرح می‌شوند، طراحی قانون هدایت با فرض سرعت نزدیک‌شوندگی ثابت و برابر با 1 km/s انجام‌شده است. همچنین، شرط شروع هدایت فاز نهایی، رسیدن ره‌گیر به چهار کیلومتری هدف در نظر

$$V_{\infty}(x(t), t) \geq \int_t^T r(x, u_{\infty}(x, \tau), \tau) d\tau + V_{\infty}(x(T), T). \quad (48)$$

با توجه به بند ۲، $\lim_{T \rightarrow \infty} V_{\infty}(x(T), T) = 0$ است. از این‌رو رابطه فوق به‌صورت زیر ساده می‌شود:

$$V_{\infty}(x(t), t) \geq \int_t^{\infty} r(x, u_{\infty}(x, \tau), \tau) d\tau. \quad (49)$$

طبق تعاریفی که در روابط (۱۱) و (۱۲) ارائه شد، واضح است که:

$$\begin{aligned} & \int_t^{\infty} r(x, u_{\infty}(x, \tau), \tau) d\tau \\ & \geq \min_{u(\cdot)} \left\{ \int_t^{\infty} r(x, u(x, \tau), \tau) d\tau \right\} = V^O(x, t). \end{aligned} \quad (50)$$

با مقایسه دو رابطه اخیر می‌توان به ناتساوی زیر رسید:

$$V_{\infty}(x, t) \geq V^O(x, t). \quad (51)$$

بنابراین بخش اول ناتساوی (۲۹) اثبات شد.

با جایگذاری $V^O(x, t)$ در معادله (۱۴) و مقداری دست‌کاری جبری می‌توان نوشت:

$$\begin{aligned} & \left(\frac{\partial V_{\infty}(x,t)}{\partial x} \right)^T (f(x,t) + g(x,t)u_{\infty}(x,t)) + \\ & \frac{\partial V_{\infty}(x,t)}{\partial t} + r(x, u_{\infty}(x,t), t) \\ &= \left(\frac{\partial V_{\infty}(x,t)}{\partial x} \right)^T (f(x,t) + g(x,t)u_{\infty}(x,t)) + \\ & \frac{\partial V_{\infty}(x,t)}{\partial \tau} + r(x, u_{\infty}(x,t), t) - \\ & \left(\left(\frac{\partial V^O(x,t)}{\partial x} \right)^T (f(x,t) + g(x,t)u^O(x,t)) + \right. \\ & \left. \frac{\partial V^O(x,t)}{\partial \tau} + r(x, u^O(x,t), t) \right) \end{aligned} \quad (52)$$

$$\begin{aligned} &= \left(\frac{\partial V_{\infty}(x,t)}{\partial x} - \frac{\partial V^O(x,t)}{\partial x} \right)^T (f(x,t) + \\ & g(x,t)u_{\infty}(x,t)) + u_{\infty}^T(x,t) R u_{\infty}(x,t) + \\ & q(x,t) - \left(\frac{\partial V^O(x,t)}{\partial x} \right)^T g(x,t) (u^O(x,t) - \\ & u_{\infty}(x,t)) - r(x, u^O(x,t), t) - \\ & u^{OT}(x,t) R u^O(x,t) - q(x,t) + \frac{\partial V_{\infty}(x,t)}{\partial t} - \end{aligned}$$

ثوابت به صورت $\varepsilon_1 = 0.0001$ ، $\varepsilon_2 = 10000$ و $\alpha = 0.000001$ تعیین شده‌اند.

سیاست‌های اولیه استفاده‌شده در اجراهای اول تا سوم الگوریتم به ترتیب در روابط (۶۰)–(۶۲) آورده شده‌اند:

$$a_{M1_1} = x(4.1262t^4 - 0.2474t^3 + 4.0121t^2x^2 + 1.4407t^2 - 0.2594tx^2 - 0.0872t + 2.6334x^2 + 4.310), \quad (60)$$

$$a_{M2_1} = x(1.1622t^4 + 1.0254t^3 + 1.1805t^2x^2 + 1.1031t^2 + 0.9876tx^2 + 0.4854t + 0.8415x^2 + 2.6226), \quad (61)$$

$$a_{M3_1} = x(0.0875t^4 + 0.0841t^3 + 0.0878t^2x^2 + 0.0999t^2 + 0.0756tx^2 + 0.0450t + 0.0596x^2 + 2.0490). \quad (62)$$

در این مقاله برای حل مسائل بهینه‌سازی مجموع مربعات از جعبه‌ابزار SOSTOOLS استفاده شده است. همچنین در تنظیمات حل این مسائل، جعبه‌ابزار Mosek به‌عنوان حل‌کننده برنامه‌ریزی‌های نیمه‌معین تعیین شده است. در اجراهای اول، دوم و سوم، الگوریتم به ترتیب پس از ۱۸، ۲۲ و ۱۱ تکرار به همگرایی رسید و در نهایت سیاست زیربینه به صورت زیر به دست آمد:

$$a_M = \begin{cases} a_{M1_{19}} & 0 \leq t \leq 2 \\ a_{M2_{23}} & 2 < t < 3, \\ a_{M3_{12}} & t \geq 3 \end{cases} \quad (63)$$

که در آن

$$a_{M1_{19}} = 0.016160t^4x - 0.226952t^3x + 0.025473t^2x^3 + 0.913210t^2x + 0.014771tx^3 - 0.723569tx + 0.138164x^3 + 5.220950x, \quad (64)$$

$$a_{M2_{23}} = 0.83567t^4x - 1.76957t^3x + 0.07327t^2x^3 + 1.05257t^2x - 0.25725tx^3 + 0.43331tx + 0.33041x^3 + 6.18639x, \quad (65)$$

$$a_{M3_{12}} = 1.7647t^4x - 9.0368t^3x + 0.3239t^2x^3 + 9.4024t^2x - 1.8710tx^3 + 5.6479tx + 2.8128x^3 + 6.6179x. \quad (66)$$

توابع ارزش اولین و آخرین تکرارهای اجراهای اول تا سوم الگوریتم به ترتیب در شکل‌های ۲ تا ۴ رسم شده‌اند. همان‌طور که در این شکل‌ها مشاهده می‌شود، تمامی توابع ارزش نشان داده‌شده مثبت هستند و توابع ارزش آخرین تکرارها به‌وضوح مقدار کمتری از اولین تکرارها دارند.

گرفته‌شده است. با این مفروضات، رابطه (۴) می‌تواند به صورت زیر نوشته شود:

$$\ddot{\lambda} = \frac{1}{r_0 - V_c t} (2V_c \dot{\lambda} - a_M + a_T) = \frac{1}{4-t} (2\dot{\lambda} - a_M + a_T). \quad (54)$$

در معادله فوق، شتاب ره‌گیر به‌عنوان ورودی کنترلی و شتاب هدف به‌عنوان اغتشاش خارجی در نظر گرفته شده‌اند.

وقتی ره‌گیر به فاصله چهار کیلومتری از هدف رسید، با اندازه‌گیری سرعت نزدیک شونده‌گی به‌وسیله جست‌وجوگر یا هر سازوکار دیگری، می‌توان زمان موردنیاز برای رسیدن ره‌گیر به هدف را به صورت زیر محاسبه کرد:

$$t_{go} = \frac{4}{V_c}. \quad (55)$$

همان‌طور که در زیر بخش ۲-۲ بیان شد، قیود استاندارد مسائل بهینه‌سازی مجموع مربعات از جنس چندجمله‌ای هستند. بنابراین، برای استفاده از الگوریتم ۲ می‌بایست از تقریب چندجمله‌ای مدل (۵۴) استفاده شود. به‌منظور دقیق‌تر شدن این تقریب‌ها می‌توان بازه $[0, t_{go}]$ را به چند بخش تقسیم و در هر بخش یک تقریب چندجمله‌ای برای تابع $\frac{1}{4-t}$ به صورت زیر به دست آورد:

$$\frac{1}{4-t} = \begin{cases} r_{inv1} & 0 \leq t \leq 2 \\ r_{inv2} & 2 < t < 3, \\ r_{inv3} & 3 \leq t \leq 4 \end{cases} \quad (56)$$

که در آن

$$r_{inv1} = 0.047302t^2 + 0.027429t + 0.256557, \quad (57)$$

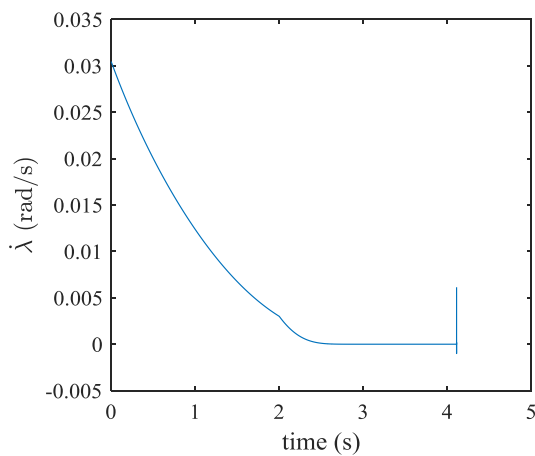
$$r_{inv2} = 0.33577t^2 - 1.17882t + 1.51404, \quad (58)$$

$$r_{inv3} = 2.8094t^2 - 16.2287t + 24.3975. \quad (59)$$

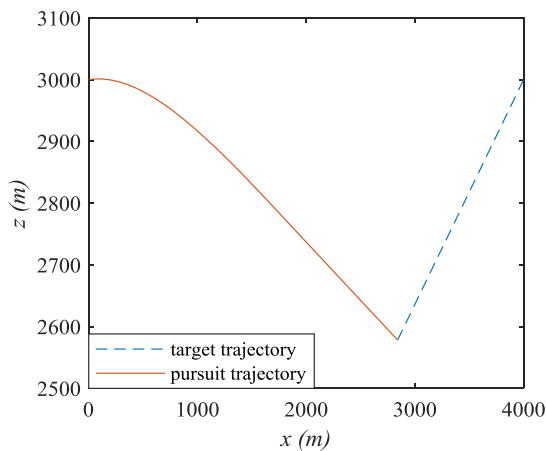
با توجه به رابطه (۵۶)، می‌بایست الگوریتم پیشنهادی سه مرتبه اجرا و هر مرتبه یکی از چندجمله‌ای‌های (۵۷)–(۵۹) به‌عنوان تقریب $\frac{1}{4-t}$ بکار گرفته شود. در ادامه، منظور از اجراهای اول، دوم و سوم الگوریتم، به ترتیب اجرای الگوریتم با در نظر گرفتن تقریب‌های چندجمله‌ای (۵۷)، (۵۸) و (۵۹) خواهد بود.

در هر سه بار اجرای الگوریتم پارامترهای طراحی از این‌قرار بوده‌اند: تابع هزینه لحظه‌ای به صورت $x^2 + 0.1u^2(x, t)$ و چندجمله‌ای‌های $P_i(x, t)$ و $V_i(x, t)$ به صورت ترکیب خطی از تک‌جمله‌ای‌های $\{x^2, x^3, tx^2, x^4, tx^3, t^2x^2\}$ و سایر

حال به‌منظور بررسی کارایی قانون هدایت طراحی‌شده، شبیه‌سازی‌هایی در قالب دو سناریوی درگیری انجام می‌شود. در سناریوی اول فرض می‌شود که در شروع فاز آشیانه‌یاب پرواز، ره‌گیر و هدف هر دو در ارتفاع 3000 m هستند. هدف با زاویه مسیر پرواز 20° و سرعت ثابت 300 m/s در حال پرواز است. ره‌گیر نیز با فاصله 4000 m و خطای نشانه‌روی -10° نسبت به هدف، با سرعت 700 m/s پرواز می‌کند. با اعمال قانون هدایت زیربینه (۶۳) به ره‌گیر و انجام شبیه‌سازی عددی نتایج نشان داده‌شده در شکل‌های ۵ و ۶ به‌دست آمده‌اند.

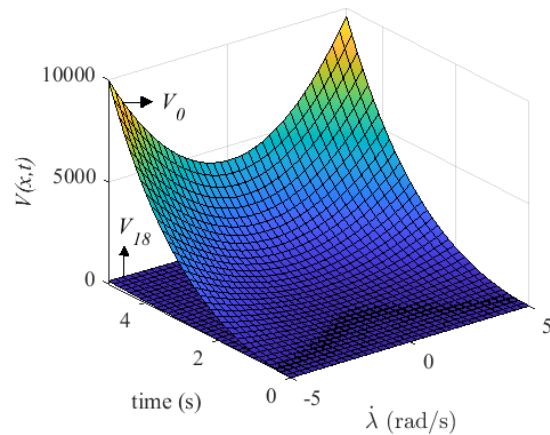


شکل (۵): نرخ چرخش خط دید در سناریوی اول.

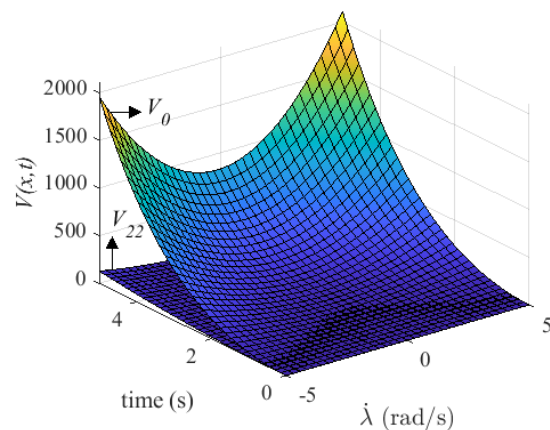


شکل (۶): مسیر برخورد ره‌گیر به هدف در سناریوی اول.

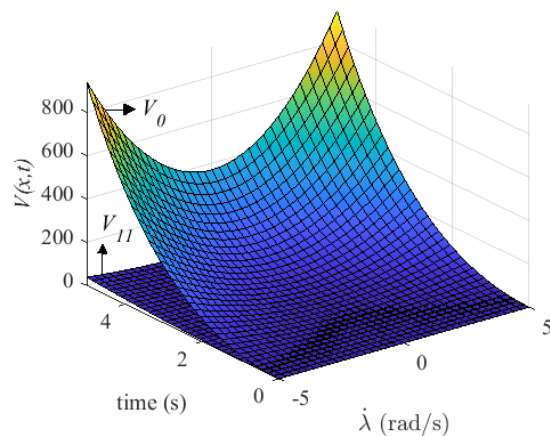
شکل ۵، نشان‌دهنده پاسخ زمانی نرخ چرخش خط دید ره‌گیر و هدف از لحظه شروع فاز نهایی پرواز است. همان‌طور که در این شکل دیده می‌شود نرخ چرخش خط دید پس از گذشت لحظاتی به صفر میل می‌کند که نتیجه آن، با توجه



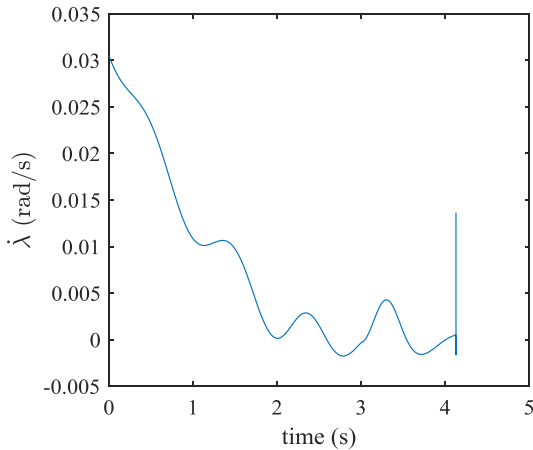
شکل (۲): توابع ارزش اولین و آخرین تکرارهای اجرای اول الگوریتم.



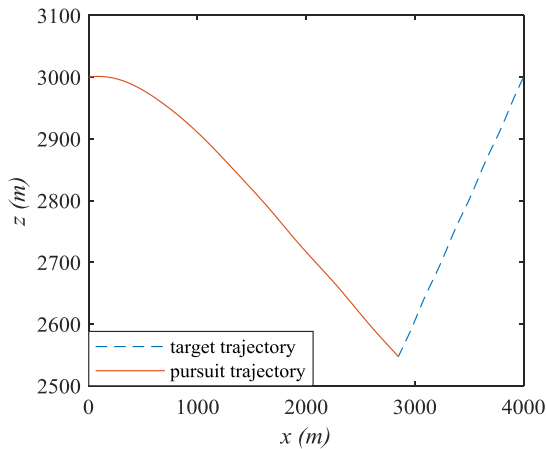
شکل (۳): توابع ارزش اولین و آخرین تکرارهای اجرای دوم الگوریتم.



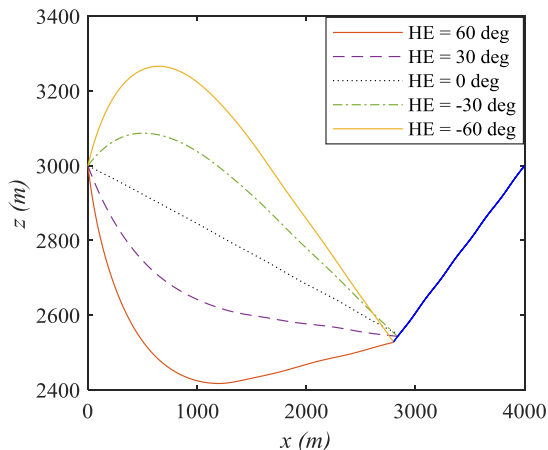
شکل (۴): توابع ارزش اولین و آخرین تکرارهای اجرای سوم الگوریتم.



شکل (۷): نرخ چرخش خط دید در سناریوی دوم.



شکل (۸): مسیر برخورد ره‌گیر به هدف در سناریوی دوم.



شکل (۹): مسیر برخورد ره‌گیر به هدف به ازای خطاهای مختلف نشانه‌روی.

به فلسفه هدایت ناوبری تناسبی، برخورد ره‌گیر به هدف می‌باشد. شکل ۶ نیز مسیر برخورد ره‌گیر به هدف در صفحه دوبعدی درگیری را نشان می‌دهد. لازم به ذکر است که از تقریب‌های چندجمله‌ای داده‌شده در رابطه (۵۶) و همچنین فرض ثابت بودن سرعت نزدیک‌شوندگی در این شبیه‌سازی‌ها استفاده‌ای نشده است و این مفروضات به‌طور خاص در بخش طراحی قانون هدایت بکار گرفته شده‌اند.

سناریوی دوم نیز همانند سناریوی اول در نظر گرفته شده است با این تفاوت که در آن سرعت هدف ثابت نیست. در این سناریو، هدف دارای شتابی به صورت $a_T = 5g \sin(2\pi t)$ فرض شده است. با اعمال قانون هدایت به ره‌گیر و انجام شبیه‌سازی، نتایج نشان‌دهنده برخورد ره‌گیر به هدف هستند. پاسخ زمانی نرخ چرخش خط دید ره‌گیر و هدف و مسیر برخورد ره‌گیر به هدف در شکل‌های ۷ و ۸ رسم شده‌اند. فاصله از دست‌دهی در این سناریو برابر با 0.4m به دست آمد که مقداری مطلوب می‌باشد.

سپس به‌منظور اثبات قوام قانون هدایت نسبت به خطاهای نشانه‌روی، سناریوی دوم با فرض خطاهای نشانه‌روی دیگری نیز مجدداً شبیه‌سازی شد. نتایج این شبیه‌سازی‌ها در شکل ۹ رسم شده‌اند. فاصله از دست‌دهی در تمامی این شبیه‌سازی‌ها مقداری کمتر از ۱ متر داشت.

درنهایت باهدف نشان دادن مزیت قانون هدایت متغیر با زمان پیشنهادی، یک قانون هدایت زیربهینه نامتغیر با زمان بر مبنای الگوریتم مبتنی بر مدل ارائه‌شده در [۱۶] طراحی و با قانون هدایت (۶۳) مقایسه شد. به دلیل محدودیت فضا از بحث در مورد جزئیات الگوریتم ارائه‌شده در [۱۶] خودداری شده است. نرخ چرخش خط دید در سناریوی دوم در صورت استفاده از روش پیشنهادی و روش مرجع [۱۶] در شکل ۱۰ رسم شده‌اند. مقدار تابع ارزش در صورت استفاده از روش پیشنهادی برابر با 0.0023 و در صورت استفاده از روش [۱۶] برابر با 0.0025 به دست آمد که نشان‌دهنده برتری روش پیشنهادی است.

angle. Journal of Guidance, Control, and Dynamics. 2008;31(5):1400-12.

[3] Nasrollahi S, Khooshehmehri A. A model of predictive terminal guidance based on whale optimization algorithm considering the aerodynamic model of the pursuer. Journal of Aerospace Mechanics. 2021;17(1):37-50.

[4] Mohammadzaman I, Momeni H. PI Guidance Law Design with Finite Time Convergence. Journal of Aerospace Mechanics. 2011;7(1).

[5] Golestani M, Mohammadzaman I, Yazdanpanah MJ, Vali AR. Application of finite-time integral sliding mode to guidance law design. Journal of Dynamic Systems, Measurement, and Control. 2015;137(11):114501.

[6] Yang C-D, Chen H-Y. Nonlinear H robust guidance law for homing missiles. Journal of Guidance, Control, and Dynamics. 1998;21(6):882-90.

[7] Kirk DE. Optimal control theory: an introduction: Courier Corporation; 2004.

[8] Bellman R. Dynamic programming. Science. 1966;153(3731):34-7.

[9] Bellman R, Dreyfus S. Functional approximations and dynamic programming. Mathematical Tables and Other Aids to Computation. 1959:247-51.

[10] Howard RA. Dynamic programming and markov processes. 1960.

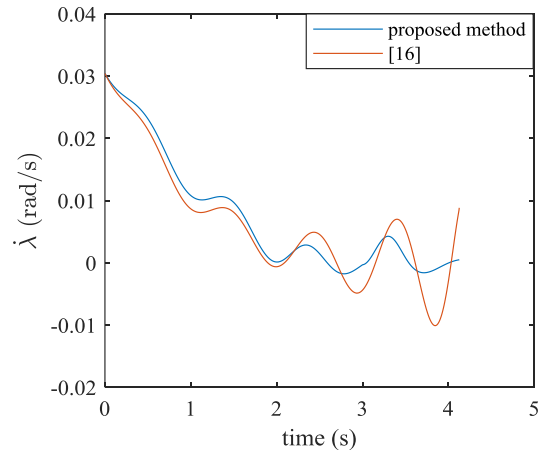
[11] Werbos P. Advanced forecasting methods for global crisis warning and models of intelligence. General System Yearbook. 1977:25-38.

[12] Kiumarsi B, Vamvoudakis KG, Modares H, Lewis FL. Optimal and autonomous control using reinforcement learning: A survey. IEEE transactions on neural networks and learning systems. 2017;29(6):2042-62.

[13] Sun J, Liu C, Ye Q. Robust differential game guidance laws design for uncertain interceptor-target engagement via adaptive dynamic programming. International Journal of Control. 2017;90(5):990-1004.

[14] Pakkhesal S, Shamaghdari S. Sum-of-squares-based policy iteration for suboptimal control of polynomial time-varying systems. Asian Journal of Control. 2021.

[15] Parrilo PA. Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization: California Institute of Technology; 2000.



شکل (۱۰): مقایسه روش پیشنهادی و روش [۱۶].

۵- نتیجه گیری

در این مقاله، مسئله طراحی قانون هدایت بهینه متغیر با زمان برای فاز آشیانه یاب رهگیرها مورد مطالعه قرار گرفت. در همین راستا، یک الگوریتم تکرار سیاست مبتنی بر بهینه سازی مجموع مربعات ارائه و از آن برای طراحی قانون هدایت استفاده شد. این الگوریتم مشکلات الگوریتم های تکرار سیاست مرسوم از جمله محلی بودن تحلیل ها، سنگین بودن محاسبات و عدم تضمین پایداری را ندارد. خاصیت هایی نیز برای الگوریتم پیشنهادی و سیاست کنترلی حاصل از آن، ادعا و اثبات شدند. این خاصیت ها شامل همگرایی الگوریتم، شدنی بودن مسئله های بهینه سازی و پایداری نمایی سیستم حلقه بسته بودند. همچنین به منظور اثبات کارایی الگوریتم پیشنهادی برای هدایت رهگیرها شبیه سازی هایی صورت گرفت و نتایج آن ها در متن مقاله گزارش شدند. از آنجایی که هدایت فاز نهایی در مدت زمان کوتاهی رخ می دهد، افزودن قابلیت تضمین پایداری زمان محدود به الگوریتم می تواند ایده خوبی برای تحقیقات بعدی در این زمینه باشد.

۶- مراجع

[1] Shneydor NA. Missile guidance and pursuit: kinematics, dynamics and control: Elsevier; 1998.

[2] Shaferman V, Shima T. Linear quadratic guidance laws for imposing a terminal intercept

- [16] Jiang Y, Jiang Z-P. Global adaptive dynamic programming for continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*. 2015;60(11):2917-29.
- [17] Zhu Y, Zhao D, He H. Invariant adaptive dynamic programming for discrete-time optimal control. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 2019;50(11):3959-71.
- [18] Yazdani NM, Moghaddam RK, Kiumarsi B, Modares H. A Safety-Certified Policy Iteration Algorithm for Control of Constrained Nonlinear Systems. *IEEE Control Systems Letters*. 2020;4(3):686-91.
- [19] Vandenberghe L, Boyd S. Semidefinite programming. *SIAM review*. 1996;38(1):49-95.
- [20] Prajna S, Papachristodoulou A, Parrilo PA. SOSTOOLS: sum of squares optimization toolbox for MATLAB-user's guide. *Control and Dynamical Systems, California Institute of Technology, Pasadena, CA*. 2004;91125.
- [21] ApS M. Mosek optimization toolbox for MATLAB. *User's Guide and Reference Manual, Version*. 2019;4.
- [22] Khalil HK. *Nonlinear systems third edition*. Patience Hall. 2002;115.



Sum-of-Squares Optimization Based Approximate Dynamic Programming for Time-Varying Systems and its Application in Suboptimal Guidance Law Design

Sajjad Pakkhesal¹, Saeed Shamaghdari^{2*}

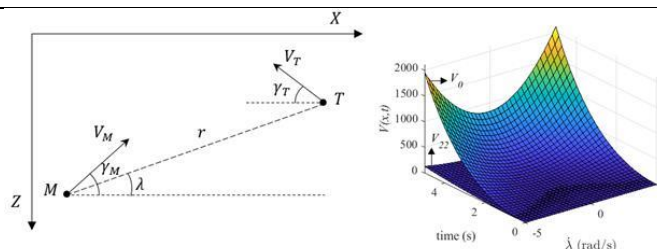
¹ Ph.D. Student, Department of Control, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.

² Associate Professor, Department of Control, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.

HIGHLIGHTS

- A policy iteration algorithm has been proposed for control of time-varying systems.
- Properties of this algorithm have been proved.
- A guidance law has been designed based on this algorithm.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Article Type: Research paper

Received: 10 July 2022

Received in revised form: 2 August 2022

Accepted: 24 August 2022

Available online: 18 October 2022

*Correspondence:

shamaghdari@iust.ac.ir

How to cite this article:

S. Pakkhesal, S. Shamaghdari. Sum-of-squares optimization based approximate dynamic programming for time-varying systems and its application in suboptimal guidance law design. Journal of Aerospace Mechanics. 2022; 18(4):89-103.

Keywords:

Sub-optimal guidance law
Approximate dynamic programming
Time-varying systems
Sum-of-squares optimization
Policy iteration algorithm

ABSTRACT

In this paper, we propose a method for sub-optimal control of time-varying polynomial systems and use it for pursuits guidance law design. Since, engagement equations between pursuit and target are depend on the range between them and this range is varying during the flight, guidance law designer is faced with a time-varying system. The developed methods for control of time-invariant systems are not directly applicable for time-varying systems. One of the conventional approaches for pursuits guidance law design is the optimal control. Approximate dynamic programming is a well-known method for solving the optimal control problem. One of the challenges of using this method for control of nonlinear time-varying systems is the difficulty of solving the Bellman equation. In the proposed method of this paper, solving the Bellman equation has been relaxed with solving a sum-of-squares optimization problem. It will be proved that the designed control policy with this method is globally exponentially stabilizing. Finally, performance of the proposed method for pursuits guidance will be illustrated with numerical simulations.

* Copyrights for this article are retained by the author(s) with publishing rights granted to Imam Hossein University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit <https://www.creativecommons.org/licenses/by-nc/4.0/legalcode>.

