

دانشگاه صنعتی شاهرود
دانشکده برق و رباتیک

پایان نامه کارشناسی ارشد برق - گرایش کنترل

عنوان

کنترل پیش بین فازی-عصبی راکتور تانک همزن پیوسته (CSTR)

محقق

اسماعیل بنی سعید

اساتید راهنما

دکتر محمد حداد ظریف

دکتر سید محمدرضا رفیعی

تشکر و قدردانی

شکر و سپاس بی‌پایان به درگاه آفریدگار متعال که لطف و عنایتش موجب شد تا تحقیق حاضر به کمک بزرگواران و عزیزانی که به گونه‌های مختلف مرا یاری دادند پایان پذیرد. از اساتید گرامی جناب آقای دکتر محمد حداد ظریف و جناب آقای دکتر سید محمدرضا رفیعی که با راهنماییها و نظارت خویش در تدوین این پایان‌نامه بمن کمک کرده‌اند سپاسگزاری نموده و موفقیت روزافزون آنان را از خداوند منان خواستارم.

همچنین از تمامی دوستانی که در مدت تحصیل در این دانشگاه به انواع مختلف باعث شده‌اند تا تحمل دشواریهای زندگی و تحصیل بدور از خانواده برای اینجانب امکانپذیر گردد تشکر و قدردانی می‌کنم. بخصوص از جناب آقای مهندس محمدمهدی جعفری کنگ بخاطر راهنماییها و حمایتهای بی‌دریغشان صمیمانه سپاسگزارم. برای تمامی این عزیزان آرزوی موفقیت، شادکامی و خوشبختی دارم.

امید است دستاوردهای این پروژه مورد استفاده دانش‌پژوهان قرار گرفته و خطاها و کاستی‌های آنرا بدیده اغماض بنگرند.

چکیده

پایان نامه حاضر تحقیقی کاربردی در مورد کنترلر پیش بین تعمیم یافته فازی-عصبی و اعمال آن به سیستم راکتور تانک همزن پیوسته می باشد. کنترلرهای پیش بین با توجه به ویژگیهای جالبی مانند بهینه سازی عملکرد سیستم ، در نظر گرفتن محدودیتهای عملی متغیرها ، شهودی بودن مفاهیم و تنوع فرایندهای قابل کنترل کاربردهای گسترده ای در کنترل فرایندهای صنعتی دارند. در این روشها به یک مدل پیش بین نیاز است که دقت این مدل بطور مستقیم بر عملکرد کنترلر تاثیرگذار می باشد. همچنین یک مساله بهینه سازی باید در هر پرپود نمونه برداری حل شود. شناسائی مدل های خطی یا استاتیک آسان بوده و مساله بهینه سازی متناظر جواب تحلیلی دارد اما اینگونه مدلها نمی توانند رفتار آینده فرایندهای غیرخطی و متغیر با زمان را با دقت مطلوب پیش بینی کنند . مدل های غیرخطی متعددی می توانند فرایندهای دینامیک را با دقت مناسب نمایش دهند اما استخراج این مدلها مشکل است و مهمتر اینکه مساله بهینه سازی در این حالت غیرخطی (و غالباً غیرمحدب) می شود که جواب تحلیلی برای آن وجود ندارد و روشهای متداول موجود برای حل آن پیچیده و وقت گیر می باشند. در این پایان نامه پس از بیان مفاهیم کلی کنترلرهای پیش بین و معرفی کنترلر پیش بین تعمیم یافته ، دو الگوریتم برای شناسائی برخط مدل غیرخطی فازی-تطبیقی تاکاگی-سوگنو ارائه شده اند. سپس با اعمال این الگوریتمها به چندین مساله مدلسازی معیار ، دقت بالا ، فشردگی ساختار و کارائی محاسباتی مدل های بدست آمده را نشان داده ایم. بدین ترتیب مشکل استخراج یک مدل غیرخطی دقیق تا حدود زیادی حل شده است. همچنین خاصیت خطی بودن زیرمدلهای تشکیل دهنده مدل غیرخطی تاکاگی-سوگنو باعث کاهش قابل توجه پیچیدگیهای حل مساله بهینه سازی در هر پرپود می شود. مدل فازی غیرخطی بدست آمده از الگوریتمهای شناسائی تطبیقی برخط مذکور در هر لحظه خطی-سازی شده و با روش کنترلر پیش بین تعمیم یافته خطی ترکیب می شود تا کنترلر پیش بین تعمیم یافته فازی-تطبیقی بدست آید. عملکرد کنترلر مذکور در کنترل راکتور تانک همزن پیوسته در شرایط مختلف (راه اندازی- تغییر پارامترها و وجود نویز اندازه گیری) مورد بررسی قرار گرفته است.

کلمات کلیدی : کنترلر پیش بین تعمیم یافته - خوشه سازی - شناسائی برخط مدل فازی تاکاگی-

سوگنو - بهینه سازی برخط - قیود عملی - کاهش بار محاسباتی - تانک راکتور همزن پیوسته

فهرست مطالب

فصل اول : مقدمه‌ای بر کنترل پیش‌بین مبتنی بر مدل

۱-۱- مقدمه	۲
۲-۱- مروری بر بعضی روشهای کنترل فرایند و لزوم استفاده از MPC	۲
۳-۱- عملیات مشترک روشهای کنترل پیش‌بین	۵
۴-۱- مزایا و معایب روشهای کنترل پیش‌بین	۹
۵-۱- پاسخ آزاد و پاسخ اجباری	۱۱
۶-۱- انواع مدل‌های پیش‌بین	۱۲
۱-۶-۱- مدل پاسخ ضربه	۱۲
۲-۶-۱- مدل پاسخ پله	۱۳
۳-۶-۱- مدل تابع انتقال	۱۴
۴-۶-۱- مدل فضای حالت	۱۴
۵-۶-۱- مدل CARIMA	۱۴
۶-۶-۱- مدل‌های غیرخطی	۱۵
۷-۱- مدل اغتشاش	۱۵

فصل دوم : کنترل پیش‌بین تعمیم‌یافته

۱-۲- مقدمه	۱۹
۲-۲- معرفی روش GPC	۱۹
۳-۲- محاسبه قانون کنترل GPC با در نظر گرفتن اغتشاشات قابل اندازه‌گیری	۲۳
۴-۲- محاسبه قانون کنترل بدون استفاده از معادله دیوفانتین	۲۶
۵-۲- انتخاب پارامترهای تابع هزینه	۲۷
۶-۲- فرمول‌بندی روش GPC در فضای حالت	۲۸
۱-۶-۲- مدل فضای حالت بدون اغتشاش	۲۸
۲-۶-۲- مدل فضای حالت با در نظر گرفتن اغتشاش	۳۱
۳-۶-۲- مدل فضای حالت به همراه رویینگر	۳۲

فصل سوم : شناسایی برخط مدل فازی TS

۱-۳- مقدمه	۳۷
------------	----

فهرست مطالب

۳۷	۲-۳- بررسی مختصر روشهای شناسائی سیستم فازی و مشکلات آنها
۴۰	۳-۳- معرفی مدل فازی TS و شبکه فازی-عصبی
۴۰	۱-۳-۳- معرفی مدل فازی TS
۴۲	۲-۳-۳- ساختار شبکه فازی-عصبی
۴۵	۴-۳- شناسائی ساختار
۴۶	۱-۴-۳- الگوریتم خوشه‌سازی تفریقی خارج خط
۴۸	۲-۴-۳- خوشه‌سازی تفریقی برخط
۵۱	۵-۳- آموزش پارامترهای زیرمدلهای خطی (شناسائی پارامتر)
۵۱	۱-۵-۳- تخمین کلی پارامترها
۵۴	۲-۵-۳- تخمین محلی پارامترها
۵۵	۶-۳- اصلاح شعاعهای تاثیر قواعد
۵۸	۷-۳- الگوریتم تکامل مدل‌های eTS
۶۵	۸-۳- شرایط تضمین پایداری الگوریتم کاهش گرادیان
۶۶	۱-۸-۳- بررسی پایداری در حالت معلوم بودن پارامترهای مقدمه
۶۷	۲-۸-۳- بررسی پایداری در حالت یادگیری پارامترهای مقدمه
۷۱	۹-۳- شبیه‌سازی
۷۱	۱-۹-۳- پیش‌بینی سری زمانی مکی-گلاس
۷۴	۲-۹-۳- پیش‌بینی داده‌های کوره گازی box-jenkins
۷۶	۳-۹-۳- شناسائی مدل CSTR
۷۹	۴-۹-۳- نتایج بدست آمده از اعمال الگوریتم به مجموعه داده‌های مختلف
۸۰	۱۰-۳- ساده‌سازی محاسباتی الگوریتم شناسائی مدل eTS

فصل چهارم : کنترل پیش‌بین تعمیم‌یافته فازی-تطبیقی

۸۷	۱-۴- مقدمه
۸۸	۲-۴- فرمول‌بندی مدل فازی TS جهت تناسب با الگوریتمهای کنترل
۸۹	۳-۴- خطی‌سازی محلی مدل TS
۸۹	۱-۳-۴- خطی‌سازی محلی تک‌مرحله‌ای مدل TS
۹۴	۲-۳-۴- خطی‌سازی محلی چندمرحله‌ای مدل TS
۹۷	۳-۳-۴- رابطه خطی‌سازی محلی با خطی‌سازی ژاکوبین

فهرست مطالب

۹۸	۴-۴- روش بهینه‌سازی $B \& B$
۱۰۳	۵-۴- الگوریتم GPC فازی-تطبیقی با استفاده از خطی‌سازی مدل TS

فصل پنجم : کنترل CSTR با روش GPC فازی

۱۰۶	۵-۱- مقدمه
۱۰۶	۵-۲- معرفی سیستم CSTR
۱۰۶	۵-۱-۲- قوانین اساسی
۱۰۹	۵-۲-۲- سینتیک شیمیائی
۱۰۹	۵-۳-۲- CSTR غیرهمدمما
۱۱۱	۵-۴-۲- مشخصات CSTR تحت کنترل
۱۱۶	۵-۳- شبیه‌سازی
۱۱۶	۵-۱-۳- ردیابی سیگنال مرجع
۱۱۸	۵-۲-۳- بررسی مقاومت الگوریتم در برابر تغییرات پارامترها
۱۱۸	۵-۱-۲-۳- تغییرات غلظت ماده A در ورودی راکتور (C_{A0})
۱۱۹	۵-۲-۲-۳- تغییرات دمای ماده A در ورودی سیستم (T_0)
۱۲۱	۵-۳-۲-۳- تغییرات همزمان دو پارامتر T_0 و C_{A0}
۱۲۲	۵-۴-۲-۳- تغییرات همزمان سه پارامتر T_{J0} و T_0 ، C_{A0}
۱۲۳	۵-۳-۳- بررسی اثر وجود نویز در اندازه‌گیری خروجی
۱۲۴	نتیجه‌گیری و پیشنهاد
۱۲۵	مراجع
۱۳۳	ضمیمه الف - اثبات برخی روابط الگوریتم شناسائی برخط مدل TS
۱۳۷	ضمیمه ب - برخی الگوریتمهای خوشه‌سازی مهم

فهرست شکل‌ها

شماره شکل	عنوان شکل	صفحه
۱-۱	وضعیت زمانی افق کنترل ، افق پیش‌بینی و افق نمونه‌برداری	۶
۲-۱	منحنی مرجع مجازی	۷
۳-۱	دیاگرام بلوکی اعمال MPC به فرایند	۹
۴-۱	مفهوم پاسخهای آزاد و اجباری	۱۱
۱-۳	نمونه‌ای از بخش‌بندی شبکه‌ای فضای دوبعدی	۴۶
۲-۳	دیاگرام بلوکی الگوریتم شناسایی مدل‌های eTS	۶۲
۳-۳	پیش‌بینی سری زمانی مکی-گلاس	۷۲
۴-۳	درصد خطا در پیش‌بینی سری زمانی مکی-گلاس با مدل فازی TS	۷۳
۵-۳	تعداد قواعد مدل فازی TS	۷۳
۶-۳	پارامترهای قسمت نتیجه قاعده سوم	۷۳
۷-۳	پارامترهای قسمت نتیجه قاعده چهارم	۷۳
۸-۳	شناسایی کوره گازی Box-Jenkins	۷۵
۹-۳	درصد خطا در شناسایی کوره گازی Box-Jenkins با مدل فازی TS	۷۵
۱۰-۳	درجه حرارت ماده A در ورودی CSTR (T_0)	۷۷
۱۱-۳	خروجی CSTR (با ورودیهای PRBS و استفاده از T_0 متغیر)	۷۷
۱-۴	دیاگرام شماتیک روش بهینه‌سازی B & B	۹۹
۲-۴	پیاده‌سازی الگوریتم GPC فازی - تطبیقی با استفاده از خطی‌سازی محلی	۹۹
۱۰۴	مدل TS به همراه طرح IMC	۱۰۴
۱-۵	دیاگرام شماتیک سیستم CSTR	۱۱۲
۲-۵	ردیابی منحنی مرجع پس از راه‌اندازی	۱۱۷
۳-۵	ردیابی منحنی مرجع در شرایط عادی	۱۱۸
۴-۵	تغییرات غلظت ماده A در ورودی راکتور و اثر آن روی خروجی	۱۱۸
۵-۵	تغییرات دمای ماده A در ورودی راکتور و اثر آن روی خروجی	۱۲۰
۶-۵	تغییرات تدریجی و مداوم دمای ماده A در ورودی و اثر آن روی خروجی	۱۲۱
۷-۵	تغییرات تدریجی و تصادفی دما و غلظت ماده A در ورودی و اثر آن روی خروجی	۱۲۱
۸-۵	تغییرات همزمان سه پارامتر و اثر آن در ردیابی سیگنال مرجع	۱۲۲
۹-۵	نویز جمع‌شونده با خروجی در حین اندازه‌گیری و اثر آن بر خروجی	۱۲۳

فهرست جداول

شماره جدول	عنوان	صفحه
۱-۳	معیارهای مورد استفاده برای اصلاح پایگاه قواعد	۶۴
۲-۳	عملکرد مدل eTS در پیش‌بینی سری مکی-گلاس (حالت D ثابت)	۷۲
۳-۳	عملکرد مدل eTS در پیش‌بینی سری مکی-گلاس (حالت D متغیر)	۷۲
۴-۳	مقایسه عملکرد الگوریتم ارائه شده با نتایج برخی مقالات	۷۴
۵-۳	عملکرد مدل eTS در پیش‌بینی داده‌های کوره گازی Box-Jenkins (D ثابت)	۷۵
۶-۳	عملکرد مدل eTS در پیش‌بینی داده‌های کوره گازی Box-Jenkins (D متغیر)	۷۵
۷-۳	نتایج شناسائی سیستم CSTR (D ثابت)	۷۶
۸-۳	نتایج شناسائی سیستم CSTR (D متغیر)	۷۷
۹-۳	نتایج شناسائی سیستم CSTR با وجود تغییرات در درجه حرارت جریان ورودی	۷۸
۱۰-۳	نتایج شناسائی سیستم CSTR (بدون استفاده از معیار حذف قواعد)	۷۸
۱۱-۳	نتایج مدل‌سازی با استفاده از الگوریتم ساده‌سازی شده	۸۴
۱۲-۳	بررسی مقاومت الگوریتم‌های مدل‌سازی در برابر نویز	۸۴
۱-۵	مقادیر پارامترهای سیستم CSTR	۱۱۵

فصل اول

مقدمه‌ای بر کنترل پیش‌بین مبتنی

بر مدل

(MPC)

۱-۱- مقدمه

در این فصل ابتدا بمنظور بیان علل گرایش به استفاده از MPC^۱، برخی از روشهای کنترل فرایند و مشکلات آنها را بطور مختصر بررسی می‌کنیم. سپس اصول کلی اعضاء خانواده MPC بیان می‌شوند و در نهایت به معرفی مدل‌های پیش‌بین متداول و بیان نقاط قوت و ضعف آنها می‌پردازیم.

۱-۲- مروری بر بعضی روشهای کنترل فرایند و لزوم استفاده از MPC

بعلت هزینه زیاد و مشکلات فنی موجود در استخراج مدل‌های دینامیکی مورد نیاز در روشهای کنترل پیشرفته و همچنین سادگی و عملکرد خوب کنترلرهای PID و عدم نیاز آنها به مدل دینامیکی، این کنترلرها هنوز کاربردهای زیادی در صنعت دارند. حتی برای سیستم‌های دارای زمان مرده قابل توجه، بخاطر وجود استعداد حساسیت به عدم تطابق مدل-سیستم در روشهای جدیدتر اغلب از همان الگوریتم PID استفاده می‌شود. تلاشهای صورت گرفته برای رفع مشکلات موجود در این زمینه به ظهور کنترل خودتنظیم^۲ و کنترل تطبیقی منجر شد. هدف از این طرحها رسیدن به کنترل مطمئن فرایندها علیرغم وجود اغتشاشات، زمان مرده و تغییرات فرایندها بوده است. اگر چه این دو روش در دهه‌های گذشته هم از جهت تئوری و هم از نظر کاربردهای عملی پیشرفتهای زیادی داشته‌اند اما هیچکدام از آنها به اندازه‌ای مناسب نبوده‌اند که بعنوان یک الگوریتم همه‌منظوره برای کنترل پایدار اکثریت فرایندهای عملی بکار گرفته شوند [۳۹]. شرط همه‌منظوره بودن روش اینست که بتوانیم آنها را به سیستم‌های زیر اعمال کنیم:

۱. فرایند غیرمی‌نیم‌فاز
- اکثر توابع انتقال پیوسته هنگامی که با سرعت کافی نمونه‌برداری شوند تمایل دارند صفرهای زمان-گسسته‌ای در بیرون دایره واحد داشته باشند.
۲. فرایند دارای ناپایداری حلقه باز و قطب‌هایی که بسختی میرا می‌شوند مانند سیستم‌های انعطاف‌پذیر.
۳. فرایند دارای زمان مرده متغیر یا نامعین

^۱ Model-based Predictive Control

^۲ Self-Tuning

بعضی روشها (مانند خودتنظیم‌های حداقل واریانس) حساسیت زیادی به فرضیات صورت گرفته در مورد زمان مرده سیستم دارند و روشهایی هم که تلاش می‌کنند تا با استفاده از اطلاعات شرایط کاری زمان مرده را تخمین بزنند، پیچیده بوده و تمایل به ناپایداری دارند.

۴. فرایند دارای مرتبه نامعین

در این حالت خودتنظیم‌های جایابی قطب و LQG عملکرد نامطلوبی خواهند داشت.

۵. فرایندهایی که ورودی-خروجی و/یا حالت‌های آنها دارای محدودیت (قید) هستند. این خصوصیت در اکثر قریب باتفاق فرایندهای واقعی وجود دارد.

روشهای قدیمی خودتنظیم مانند روشهای حداقل واریانس^۱ (MV) (آستروم^۲ و ویتن‌مارک^۳، ۱۹۷۳)؛ حداقل واریانس تعمیم یافته^۴ (GMV) (کلارک^۵، ۱۹۷۵) و جایابی قطب حقیقتاً از PID برتر می‌باشند اما این برتری تنها زمانی وجود دارد که پارامترهای کلیدی مانند مرتبه مدل و زمان مرده بدقت تعیین شده باشند. خودتنظیم MV در مقابل فرضیات صورت گرفته در مورد مرتبه مدل مقاوم است اما تجربیات عملی نشان داده‌اند که عملکرد روش MV (و همچنین روش GMV بمیزان کمتر) به انتخاب اولیه زمان مرده k (که بر اساس آن پیش‌بینی k مرحله جلوتر انجام می‌شود) وابسته است و در صورت انتخاب اشتباه زمان مرده یا تغییر آن ممکن است ناپایدار شود. از طرف دیگر روش جایابی قطب می‌تواند برای سیستمهای دارای زمان مرده متغیر بکار برده شود (بشرطی که دینامیک صورت بطور مناسب پارامتری شده باشد) ولی در حالت وجود عوامل مشترک در چندجمله‌ایهای مدل تخمینی ممکن است این روش نیز با شکست مواجه شود.

از حوالی سال ۱۹۷۰ تاکنون تعدادی روش کنترل پیش‌بین دامنه وسیع^۶ (LRPC) طراحی شده‌اند که با داشتن برخی جنبه‌های جذاب تا حدود زیادی شرایط فوق‌الذکر را برآورده می‌کنند و با بهینه‌سازی

^۱ Minimum Variance

^۲ Astrom

^۳ Wittenmark

^۴ Generalized Minimum Variance

^۵ Clarke

^۶ Long Range Predictive Control

عملکرد سیستم هزینه‌ها را کاهش می‌دهند. همچنین بعلا در نظر گرفتن قیود عملی متغیرها در حل مساله بهینه‌سازی، پیاده‌سازی سخت‌افزاری سیگنال‌های کنترل محاسبه شده مشکل نخواهد بود. بنا بدلائل فوق از MPC بعنوان یک الگوی کنترل استاندارد در زمینه فرایندهای صنعتی نام برده می‌شود. از بین روشهای خانواده LRPC می‌توان به موارد زیر اشاره کرد: روش^۱ MAC (ریچالت و دیگران، ۱۹۷۸)، روش^۲ DMC (کاتلر و راماکر، ۱۹۸۰)، روش^۳ EPSAC (دیکیزر و ونکاونبرگ، ۱۹۸۱)، روش^۴ EHAC، روش^۵ GPC (کلارک و دیگران، ۱۹۸۷)، روش^۶ CRHPC (کلارک و اسکاتولینی، ۱۹۹۱)، روش^۷ SGPC (کوواریتاکیس و دیگران، ۱۹۹۲)، روش^۸ IHGPC، روش^۹ MMPC (یو، ۱۹۹۷)؛ کاساولا و دیگران، ۲۰۰۰).

روش EPSAC از تابع انتقال گسسته (تبدیل Z) برای مدلسازی فرایند استفاده می‌کند. ساختار قانون کنترل آن ساده بوده و بصورت تحلیلی با استفاده از تابع عملکرد درجه دوم محاسبه می‌شود (با فرض اینکه $u(t)$ از لحظه t به بعد ثابت می‌ماند). مدل فرایند می‌تواند اغتشاشات قابل اندازه‌گیری را نیز دربرگیرد.

روش GPC از تابع عملکرد درجه دوم^{۱۰} و ضریب وزنی برای انرژی سیگنال کنترل استفاده می‌کند و روشی برای حل مساله بهینه‌سازی ارائه می‌دهد. روشهای GPC و DMC عمومیت بیشتری دارند. روش CSGPC (GPC مقید پایدار) قابلیت اجرا و پایداری روش GPC را در شرایط وجود قیود محدودکننده نیز دارد.

^۱ Model Algorithmic Control

^۲ Dynamic Matrix Control

^۳ Extended Prediction Self-Adaptive Control

^۴ Extended Horizon Adaptive Control

^۵ Generalized Predictive Control

^۶ Constrained Receding-Horizon Predictive Control

^۷ Stable Generalized Predictive Control

^۸ Infinite-Horizon Generalized Predictive Control

^۹ Min-Max Predictive Control

^{۱۰} Quadratic

۱-۳- عملیات مشترک روشهای کنترل پیش‌بین

- با توجه به شکل (۱-۱) در فاصله زمانی بین هر دو نمونه‌گیری متوالی، عملیات زیر انجام می‌شوند:
- ۱- در هر لحظه زمانی k با استفاده از مدل دینامیکی فرایند، خروجی $y(k+j)$ در بازه زمانی $j = H_{1P}, \dots, H_{2P}$ پیش‌بینی می‌شود. خروجی پیش‌بینی شده در لحظه k ام را با $\hat{y}(k+j|k)$ نشان می‌دهیم. این پیش‌بینی به ورودیها و خروجیهای گذشته و همچنین رشته کنترل آینده $(u(k+j|k), j=0,1,\dots,H_C)$ بستگی دارد. مقدار H_P را افق پیش‌بینی^۱ می‌نامیم. H_{1P} و H_{2P} حداقل و حداکثر افق پیش‌بینی هستند و افق کنترل نیز با H_C نمایش داده شده است.
 - ۲- خروجی واقعی اندازه‌گیری شده و تفاوت بین خروجی پیش‌بینی شده و خروجی مطلوب محاسبه می‌شود (خطای ردیابی پیش‌بینی شده).
 - ۳- رشته کنترلی بهینه $(u(k+j|k), j=0,1,\dots,H_C)$ بر اساس اندازه‌گیریها و خطای ردیابی پیش‌بینی شده بگونه‌ای محاسبه می‌شود که تابع هزینه حداقل شود. تابع هزینه نوعی مورد استفاده در روشهای MPC بصورت رابطه (۱-۱) می‌باشد.

$$J(H_{1P}, H_{2P}, H_C) = \sum_{j=H_{1P}}^{H_{2P}} \alpha_i [y_{ref}(k+j) - \hat{y}(k+j|k)] + \sum_{j=1}^{H_C} \beta_i \Delta u(k+j-1)^2 \quad (1-1)$$

پارامترهای α_i ، β_i ، y_{ref} به ترتیب سیگنال مرجع (یا تقریبی از آن)، ضریب وزنی خطای پیش‌بینی و ضریب وزنی انرژی سیگنال کنترل می‌باشند. در بعضی روشها جمله دوم تابع هدف حذف می‌شود و بعضی دیگر بجای تغییرات سیگنال کنترل از خود سیگنال در تابع هدف استفاده می‌کنند.

تغییرات سیگنال کنترل بعد از بازه H_C ($H_C < H_{2P}$) صفر فرض می‌شود:

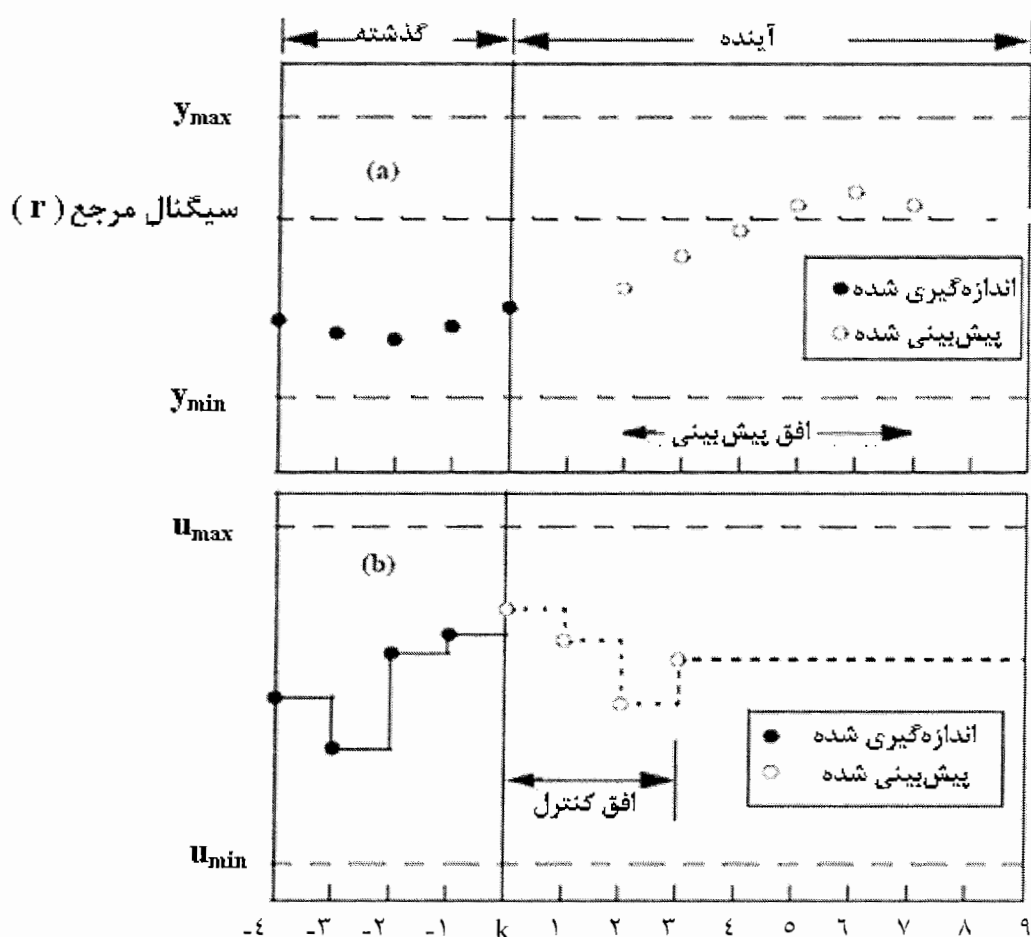
$$\Delta u(t+j-1) = 0 \quad ; j > H_C \quad (2-1)$$

این تصمیم تنها جهت سادگی محاسبات اتخاذ شده و محدودیتی ایجاد نمی‌کند. برای بیان علت این امر فرض کنیم در لحظه حاضر بین خروجی واقعی و خروجی مطلوب خطا وجود دارد. افق کنترل باید بگونه‌ای انتخاب شود که مطمئن شویم بعد از سپری شدن زمانی باندازه افق کنترل (یعنی بعد از لحظه

^۱ Prediction Horizon

خروجی واقعی باندازه کافی به خروجی مطلوب نزدیک شده است بصورتی که بعد از افق کنترل دیگر نیازی به تغییر سیگنال کنترل و افزایش بی‌رویه حجم محاسبات وجود نداشته باشد.

۴- اعمال/استراتژی/افق پس‌رونده^۱، بدین معنی که اولین عنصر از رشته کنترل محاسبه شده $u(k|k)$ به سیستم واقعی اعمال می‌شود و از بقیه عناصر رشته کنترل صرف‌نظر می‌شود، اندازه‌گیریها مجدداً انجام شده و الگوریتم به مرحله اول برمی‌گردد. (یعنی در فاصله زمانی بین هر دو نمونه‌گیری متوالی تمام مراحل پیش‌بینی، بهینه‌سازی و اعمال کنترل باید انجام شوند).



شکل ۱-۱- وضعیت زمانی افق کنترل، افق پیش‌بینی و لحظات نمونه‌برداری

^۱ Receding Horizon

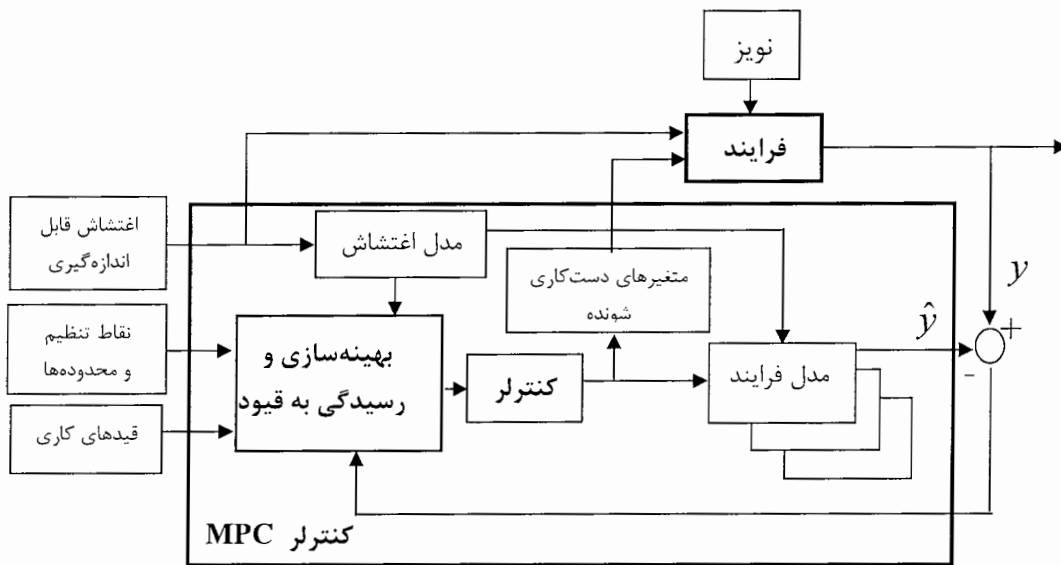
تاخیرهای زمانی فرایند یا کندی پاسخ آن تا حدود زیادی جبران می‌شود. فاصله زمانی بین تحت تاثیر قرار گرفتن پاسخ سیستم در اثر تغییر قریب‌الوقوع در منحنی مرجع تا ایجاد تغییر واقعی در سیگنال مرجع به افقهای پیش‌بینی و کنترل بستگی دارد.

ممکن است این سوال مطرح شود که چرا تمام رشته کنترلی محاسبه شده به سیستم واقعی اعمال نمی‌شود. می‌دانیم که در محاسبه سیگنالهای کنترل از پیش‌بینی‌های بدست آمده از مدل فرایند، مدل اغتشاش قابل اندازه‌گیری و وضعیت آینده منحنی مرجع استفاده می‌شود. در مورد مدل‌های فرایند و اغتشاش قابل اندازه‌گیری با توجه به مشکلات موجود در مدل‌سازی، خاصیت تصادفی اغتشاش و وجود نویز غیرقابل اندازه‌گیری، همیشه مقداری تفاوت بین خروجی آینده واقعی و خروجی آینده پیش‌بینی شده وجود دارد. از طرف دیگر در برخی کاربردها مقادیر آینده منحنی مرجع معلوم نبوده یا ممکن است با مقداری که در لحظه حاضر در بهینه‌سازی بکار رفته تفاوت پیدا کنند. بنابراین نباید به تغییرات غیرقابل پیش‌بینی که ممکن است در طول افق پیش‌بینی اتفاق افتند بی‌توجه باشیم بلکه باید با توجه به خروجی اندازه‌گیری شده در هر لحظه و ورودی-خروجیهای قبلی، پیش‌بینی‌های ذکر شده در هر لحظه تکرار شوند و چون این پیش‌بینی‌ها بر جواب مساله بهینه‌سازی (سیگنال کنترل) تاثیرگذار هستند بنابراین در هر لحظه تنها اولین عنصر از رشته کنترلی را به سیستم اعمال کرده و مساله بهینه‌سازی را دوباره با استفاده از پیش‌بینی‌های جدید حل می‌کنیم.

شکل (۱-۳) نحوه اعمال کنترلر MPC را نشان می‌دهد. منظور از "متغیرهای دستکاری شونده"^۱ در حقیقت همان اولین عنصر از رشته کنترلی بهینه محاسبه شده است. قیود کاری و نقاط تنظیمی^۲ به کنترلر داده می‌شوند و جهت جبران پیش‌خور اغتشاشات، مدلی برای اغتشاشات قابل اندازه‌گیری در نظر گرفته می‌شود.

^۱ Manipulated Variables

^۲ Set point



شکل ۳-۱ - دیاگرام بلوکی اعمال MPC به فرایند

۴-۱- مزایا و معایب روشهای کنترل پیش‌بین

خانواده MPC نسبت به دیگر روشهای کنترل دارای مزایای زیر می‌باشند :

۱. مدل دینامیکی فرایند بطور مستقیم در محاسبه سیگنال کنترل دخالت دارد که این ویژگی به کنترلر اجازه می‌دهد تا در هر لحظه با همه جنبه‌های قابل توجه دینامیک سیستم ارتباط داشته باشد .
۲. اعضای خانواده کنترل پیش‌بین برای کاربرانی که دانش کنترلی آنها محدود است جذابیت زیادی دارند چون مفاهیم آنها ملموس بوده و در عین حال انجام تنظیمات آنها نسبتاً آسان می‌باشد.
۳. فرایندهای بسیار متنوعی را می‌توان با استفاده از روشهای MPC کنترل کرد : از سیستمهای دارای دینامیک ساده تا فرایندهای پیچیده‌تر شامل سیستمهای دارای تاخیر زمانی (زمان مرده) ، فرایندهای غیرمی‌نیمم فاز ، ناپایدار حلقه‌باز و چندمتغیره.
۴. این روشها خاصیت جبران زمان مرده و همچنین جبران خطای سنسور یا عملگر را بصورت ذاتی دارند [۳۱].

۵. الگوریتم‌های MPC رفتار آینده فرایند در افق پیش‌بینی را در نظر می‌گیرند بنابراین آثار اغتشاشات قابل اندازه‌گیری را می‌توان پیش‌بینی و حذف کرد.

۶. در این روشها قیود موجود روی ورودیها، حالات سیستم و خروجیها مستقیماً در محاسبه سیگنال کنترل در نظر گرفته می‌شوند بنابراین سیگنال کنترل بدست آمده با ملاحظات عملی سازگاری داشته و پیاده‌سازی سخت‌افزاری آن آسان می‌باشد.

از طرف دیگر تکنولوژی صنعتی MPC موجود دارای محدودیتهای گوناگونی است :

۱- اکثر محصولات تجاری از مدل‌های پاسخ ضربه یا پله استفاده می‌کنند که تعداد پارامترهای زیادی دارند. این مشکل با کاربرد مدل پارامتری اتورگرسیو حل می‌شود.

۲- قواعد انتخاب پارامترهای روشهای MPC مانند افقهای پیش‌بینی و کنترل تنها یک راهنمای کلی برای اینکار می‌باشند و نمی‌توانند این پارامترها را برای هر فرایند بصورت دقیق مشخص کنند. همچنین نحوه برقراری مصالحه بین پارامترها و رفتار حلقه بسته سیستم بطور دقیق مشخص نیست.

۳- اگر چه بسته‌های شناسائی مدل می‌توانند عدم قطعیت‌های مدل را تخمین بزنند اما تنها تعداد اندکی از آنها از این اطلاعات برای طراحی کنترلر استفاده می‌کند. اکثر مدل‌های متداول بصورت خارج خط بدست آمده و نمی‌توانند سیستمهای دینامیکی و متغیر با زمان را بخوبی نمایش دهند.

۴- در فرمول‌بندیهای اولیه MPC با افق متناهی، تحلیل سیستماتیک خصوصیات پایداری و مقاومت امکانپذیر نیست. قانون کنترل متغیر با زمان بوده و بخصوص در حالت مقید، امکان نمایش آن بفرم حلقه بسته استاندارد وجود ندارد.

اسامی چند مورد از کنترلرهای پیش‌بین که بصورت تجاری ارائه شده‌اند از این قرار می‌باشند :

- **AspenTech** : Dynamic Matrix Control
- **Adersa** : Identificatin and Command (IDCOM), Hierarchical Constraint Control (HIECON), and Predictive Functional Control
- **Honeywell Profimatics** : Robust Model Predictive Control Technonlogy (RMPCT) and Predictive Control Technology(PCT)
- **Setpoint Inc** : Setpoint Multivariable Control Architecture (SMCA) and (IDCOM)-M(multivariable)

- **ABB: DMPC**

لازم به ذکر است که هر کدام از این محصولات لزوماً یک روش کنترل جداگانه نیست بلکه بسته‌هایی که اغلب شامل تست‌های شناسایی و غیره هستند نیز به هر کدام اضافه شده‌اند.

۵-۱- پاسخ آزاد و پاسخ اجباری

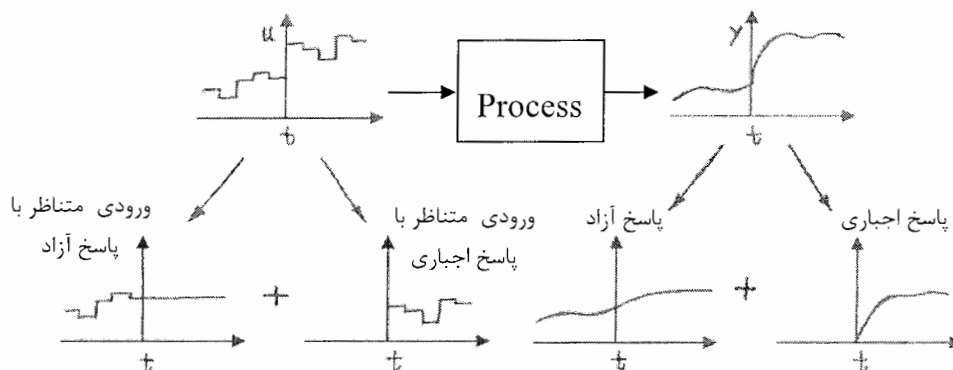
ویژگی نوعی اکثر روش‌های MPC استفاده از مفاهیم پاسخ‌های آزاد و اجباری می‌باشد. ایده اینست که رشته کنترل بصورت مجموع دو سیگنال بیان شود. $(u(t) = u_f(t) + u_c(t))$. سیگنال $u_f(t)$ (ورودی متناظر با پاسخ آزاد) ورودیهای گذشته را بیان می‌کند و در لحظات آینده بصورت ثابت و برابر با آخرین مقدار متغیر دستکاری‌شونده در نظر گرفته می‌شود:

$$\begin{aligned} u_f(t-j) &= u(t-j) \quad \text{for } j=1,2,\dots \\ u_f(t+j) &= u(t-1) \quad \text{for } j=0,1,2,\dots \end{aligned} \quad (3-1)$$

سیگنال $u_c(t)$ (ورودی متناظر با پاسخ اجباری) در لحظات گذشته برابر صفر فرض می‌شود و در لحظات آینده برابر با تغییرات سیگنال کنترل خواهد بود :

$$\begin{aligned} u_c(t-j) &= 0 \quad \text{for } j=1,2,\dots \\ u_c(t+j) &= u(t+j) - u(t-1) \quad \text{for } j=0,1,2,\dots \end{aligned} \quad (4-1)$$

شکل (۴-۱) مفهوم پاسخ‌های آزاد و اجباری را بیان می‌کند.



شکل ۴-۱- مفهوم پاسخ‌های آزاد و اجباری

۱-۶- انواع مدل‌های پیش‌بین

مدل پیش‌بین باید بتواند دینامیک فرایند را بخوبی نمایش داده و رفتار آینده آنرا بدقت پیش‌بینی کند. در عین حال مدل باید شهودی بوده و امکان بررسی تئوری مساله را فراهم کند. استراتژی‌های مختلف MPC از مدل‌های گوناگونی استفاده کرده‌اند. ورودی‌های قابل اندازه‌گیری بدو دسته ورودی‌های دستکاری-شونده^۱ و اغتشاش‌های قابل اندازه‌گیری تقسیم می‌شوند. همچنین مدل اغتشاش جهت در نظر گرفتن رفتارهایی بکار می‌رود که توسط مدل فرایند انعکاس داده نمی‌شوند (مانند ورودی‌های غیرقابل اندازه‌گیری، نویز و خطای مدل). می‌توانیم مدل کلی بدو قسمت مدل فرایند و مدل اغتشاش تقسیم کنیم. این بخش به معرفی مدل‌های پرکاربرد اختصاص دارد.

۱-۶-۱- مدل پاسخ ضربه

در این مدل که با نام‌های رشته وزنی^۲ یا مدل کانولوشن نیز شناخته می‌شود داریم:

$$y(t) = \sum_{i=1}^{\infty} h_i u(t-i) \quad (5-1)$$

که h_i نمونه‌های خروجی به ازاء ورودی ضربه واحد می‌باشد. مجموع بالا بطور کامل محاسبه نمی‌شود بلکه تنها N جمله اول آن در نظر گرفته می‌شوند. بنابراین مدل پاسخ ضربه تنها برای فرایندهای پایدار بدون انتگرالگیر استفاده می‌شود. در این حالت خواهیم داشت:

$$y(t) = \sum_{i=1}^N h_i u(t-i) = H(z^{-1})u(t) \quad (6-1)$$

که داریم:

$$H(z^{-1}) = h_1 z^{-1} + h_2 z^{-2} + \dots + h_N z^{-N} \quad (7-1)$$

عیب دیگر مدل پاسخ ضربه زیاد بودن پارامترهای مورد نیاز می‌باشد چون N غالباً مقداری بزرگ (در محدوده ۴۰ الی ۵۰) می‌باشد. معادله پیش‌بین با استفاده از مدل پاسخ ضربه باینصورت بدست می‌آید:

^۱ Manipulated Variables

^۲ Weightig sequence

$$\hat{y}(t+k|t) = \sum_{i=1}^N h_i u(t+k-i|t) = H(z^{-1})u(t+k|t) \quad (8-1)$$

این مدل شهودی بوده و اثر هر متغیر دستکاری‌شونده روی خروجی را بوضوح نشان می‌دهد. همچنین براحتی می‌توان آنرا در مورد سیستم‌های چندمتغیره بکار برد. در حالت خواهیم MIMO داشت :

$$y_j(t) = \sum_{k=1}^m \sum_{i=1}^N h_i^{k,j} u^k(t-i) \quad (9-1)$$

یک مزیت بزرگ پاسخ ضربه اینست که هیچگونه نیازی به اطلاعات اولیه فرایند ندارد. بنابراین فرایند شناسائی ساده می‌شود و درعین حال مدل بدست آمده براحتی می‌تواند دینامیکهای پیچیده‌تری نظیر خواص غیرمی‌نیمم فاز یا تاخیر زمانی را براحتی نمایش دهد. مدل پاسخ ضربه در روش MAC و (بعنوان یک حالت خاص) در روشهای GPC و ESPAC کاربرد دارد.

۱-۶-۲- مدل پاسخ پله

این مدل شباهت زیادی به مدل پاسخ ضربه دارد جز اینکه برای بدست آوردن ضرایب آن ورودی سیستم بصورت پله انتخاب می‌شود. خروجی کوتاه‌شده^۱ این مدل بشکل زیر بیان می‌شود :

$$y(t) = y_0 + \sum_{i=1}^N g_i \Delta u(t-i) = y_0 + G(z^{-1})(1-z^{-1}) \quad (12-1)$$

که g_i مقادیر نمونه‌برداری شده خروجی به ازاء ورودی پله می‌باشند. بدون از دست دادن کلیت مطلب می‌توانیم y_0 را برابر صفر فرض کنیم. در اینصورت مدل پیش‌بین باینصورت خواهد بود :

$$\hat{y}(t+k|t) = \sum_{i=1}^N g_i \Delta u(t+k-i|t) \quad (11-1)$$

مزایا و معایب مدل پاسخ پله همانند مدل پاسخ ضربه می‌باشند.

^۱ Truncated Output

۱-۶-۳- مدل تابع انتقال

این مدل در روشهای GPC، UPC، EPSAC، EHAC بکار رفته است. خروجی مدل و معادله پیش‌بین با استفاده از رابطه زیر بدست می‌آید:

$$A(z^{-1})y(t) = B(z^{-1})u(t) \implies \hat{y}(t+k|t) = \frac{B(z^{-1})}{A(z^{-1})}u(t+k|t) \quad (12-1)$$

مدل تابع انتقال می‌تواند فرایندهای ناپایدار را نیز نمایش دهد و مزیت آن کم بودن تعداد پارامترهای مورد نیاز می‌باشد. قابلیت فهم این مدل در صنعت از دیگر مدلها بیشتر می‌باشد چون مفاهیمی مانند زمان مرده، بهره و ثابتهای زمانی که معمولاً در صنعت مورد نیاز هستند بوضوح در نمایش تابع تبدیل قابل تشخیص هستند اما از طرف طراحی کنترلر برای این مدل مشکلتر می‌باشد.

۱-۶-۴- مدل فضای حالت

این مدل برای بررسی تئوری مناسبتر است و طراحی کنترلر برای آن (حتی در حالت چندمتغیره) آسان می‌باشد. برای پیش‌بینی با استفاده از این مدل داریم:

$$\hat{y}(t+k|t) = C \hat{x}(t+k|t) = C \left[A^k x(t) + \sum_{i=1}^k A^{i-1} B u(t+k-i|t) \right] \quad (13-1)$$

با استفاده از این مدل قانون کنترل از فیدبک کردن ترکیب خطی متغیرهای حالت بدست می‌آید.

۱-۶-۵- مدل CARIMA

تعداد زیادی از روشهای کنترل پیش‌بین مبتنی بر مدل خطی فرایند می‌باشند. هنگامیکه مساله تنظیم فرایند حول یک نقطه کار مد نظر باشد می‌توان مدل خطی شده^۱ CARMA را بکار برد (حتی برای فرایندهای غیرخطی)

$$A(z^{-1})y(t) = z^{-d}B(z^{-1})u(t-1) + C(z^{-1})\xi(t) \quad (14-1)$$

^۱ Controlled Auto-Regressive And Moving-Average

که در این رابطه داریم :

$$A(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_{n_a} z^{-n_a}$$

$$B(z^{-1}) = b_0 + b_1 z^{-1} + \dots + b_{n_b} z^{-n_b}$$

$$C(z^{-1}) = 1 + c_1 z^{-1} + \dots + c_{n_c} z^{-n_c}$$

در روابط بالا ، z^{-1} عملگر تاخیر باندازه یک پله زمانی ، d تاخیر زمانی خالص سیستم و $\xi(t)$ رشته تصادفی غیرهمبسته^۱ (نویز سفید با میانگین صفر) می‌باشد. مدل مذکور برای کاربردهای صنعتی دارای اغتشاشات غیر ایستا^۲ نامناسب بوده و بجای آن از مدل زیر استفاده می‌شود :

$$A(z^{-1})y(t) = z^{-d}B(z^{-1})u(t-1) + C(z^{-1})\xi(t)/\Delta \quad (۱۵-۱)$$

که Δ ، عملگر تفاضل $1 - z^{-1}$ است . مدل فوق^۳ CARIMA نامیده می‌شود.

۱-۶-۶- مدلهای غیرخطی

تواناییهای مدلهای غیرخطی مانند مدلهای مبتنی بر شبکه‌های عصبی ، مدلهای فازی و غیره برای نمایش دقیق دینامیک‌های فرایند از مدلهای خطی و استاتیک بیشتر است اما از طرف دیگر این استخراج این مدلها مشکل‌تر می‌باشد. همچنین مدلهای غیرخطی باعث افزایش پیچیدگی مساله بهینه‌سازی می‌شوند.

۱-۷- مدل اغتشاش

وجود اغتشاش باعث ایجاد تفاوت بین خروجی اندازه‌گیری شده و خروجی واقعی شده و اثرات منفی دیگری نیز دارد بنابراین اهمیت مدل اغتشاش از مدل فرایند کمتر نیست . در مدل CARIMA اغتشاش بصورت زیر نمایش داده می‌شود:

^۱ Uncorrelated

^۲ Nonstationary

^۳ Controlled Auto-Regressive And Integrated Moving-Average

$$n(t) = \frac{C(z^{-1})}{D(z^{-1})} e(t) \quad (16-1)$$

چندجمله‌ای $D(z^{-1})$ شامل انتگرالگیر $\Delta = 1 - z^{-1}$ است و $e(t)$ نویز سفید با میانگین صفر است. چندجمله‌ای C نیز عموماً بصورت واحد در نظر گرفته می‌شود.

مدل CARIMA در روشهای GPC، EPSAC و EHAC بکار گرفته می‌شود و با اندکی تغییرات در الگوریتمهای دیگر نیز استفاده می‌شود. توجه شود که در نظر گرفتن انتگرالگیر باعث دستیابی به کنترل حالت پایدار "آزاد از آفست"^۱ می‌شود.

اکنون معادله دیوفانتین را بصورت زیر معرفی می‌کنیم:

$$1 = E_k(z^{-1})D(z^{-1}) + z^{-k}F_k(z^{-1}) \quad (17-1)$$

چندجمله‌ایهای E_j و F_j بصورت یکتا بترتیب با درجه‌های $(j-1)$ و n_a تعریف می‌شوند. برای بدست آوردن این چندجمله‌ایها، عدد ۱ را آنقدر بر چندجمله‌ای $D(z^{-1})$ تقسیم می‌کنیم تا باقیمانده را بتوان بصورت $F_j(z^{-1})z^{-j}$ فاکتورگیری کرد. خارج‌قسمت تقسیم همان E_j خواهد بود. با استفاده از معادله دیوفانتین داریم:

$$n(t) = E_k(z^{-1})e(t) + z^{-k} \frac{F_k(z^{-1})}{D(z^{-1})} e(t) \quad (18-1)$$

$$n(t+k) = E_k(z^{-1})e(t+k) + F_k(z^{-1})n(t) \quad (19-1)$$

بنابراین بعنوان معادله پیش‌بین اغتشاش خواهیم داشت:

$$\hat{n}(t+k|t) = F_k(z^{-1})n(t) \quad (20-1)$$

اگر معادله بالا با مدل تابع انتقال ترکیب شود آنگاه $D(z^{-1}) = A(z^{-1})(1-z^{-1})$ ، در نتیجه داریم:

^۱ Offset free

$$\hat{y}(t+k|t) = \frac{B(z^{-1})}{A(z^{-1})}u(t+k|t) + F_k(z^{-1}) \left(y(t) - \frac{B(z^{-1})}{A(z^{-1})}u(t) \right) \quad (21-1)$$

$$\hat{y}(t+k|t) = F_k(z^{-1})y(t) + \frac{B(z^{-1})}{A(z^{-1})} \left(1 - z^{-k} F_k(z^{-1}) \right) u(t+k|t) \quad (22-1)$$

سپس با استفاده از معادله دیوفانتین خواهیم داشت :

$$\hat{y}(t+k|t) = F_k(z^{-1})y(t) + E_k(z^{-1})B(z^{-1})\Delta u(t+k|t) \quad (23-1)$$

در یک حالت خاص مدل ARIMA، مدل اغتشاش بصورت زیر در نظر گرفته می‌شود:

$$n(t) = \frac{e(t)}{(1-z^{-1})} \quad (24-1)$$

بنابراین بهترین پیش‌بینی اغتشاش بصورت $\hat{n}(t+k|t) = n(t)$ خواهد بود. این مدل اغتشاش به‌مراه مدل پاسخ پله در الگوریتم DMC و روشهای مرتبط با آن استفاده می‌شود.

فصل دوم

کنترل پیش بین تعمیم یافته (GPC)

۲-۱- مقدمه

در این فصل به معرفی یکی از مهمترین الگوریتم‌های خانواده کنترل پیش‌بین (کنترل پیش‌بین تعمیم‌یافته) می‌پردازیم. در ابتدا این روش را با استفاده از مدل CARIMA و معادله دیوفانتین فرمول‌بندی می‌کنیم. این فرمول‌بندی بدلیل استفاده صریح از مفهوم پاسخ آزاد و اجباری ملموس است اما استفاده از آن در حالت چندمتغیره مشکل می‌باشد. با توجه به مزایای مدل فضای حالت در طراحی کنترلر و کار با سیستم‌های چندمتغیره، بقیه فصل به بیان فرمول‌بندیهای GPC در فضای حالت اختصاص یافته است.

۲-۲- معرفی روش GPC

در روش GPC هدف محاسبه رشته سیگنال کنترلی آینده $u(t), \dots, u(t+H_C)$ بگونه‌ای است که خروجی در لحظات آینده تا حد امکان به منحنی مطلوب نزدیک شود. اینکار با حداقل‌سازی تابع هدف (۱-۱) انجام می‌شود. بدین منظور باید مقادیر آینده خروجی برای $H_{1P} \leq j \leq H_{2P}$ تخمین زده شوند $(\hat{y}(t+j|t))$. معادله دیوفانتین را بصورت زیر در نظر بگیرید:

$$1 = E_j(z^{-1})\tilde{A}(z^{-1}) + z^{-j}F_j(z^{-1}) \quad (1-2)$$

که در این معادله داریم:

$$\tilde{A}(z^{-1}) = (1 - z^{-1})A(z^{-1}) \quad (2-2)$$

با ضرب رابطه مدل خطی CARIMA فرایند (معادله ۱-۱۵) در $E_j(z^{-1})z^j$ خواهیم داشت:

$$\tilde{A}(z^{-1})E_j(z^{-1})y(t+j) = E_j(z^{-1})B(z^{-1})\Delta u(t+j) + E_j(z^{-1})e(t+j) \quad (3-2)$$

و از ترکیب معادلات (۲-۲) و (۱-۲) به رابطه زیر می‌رسیم:

$$(1 - z^{-j}F_j(z^{-1}))y(t+j) = E_j(z^{-1})B(z^{-1})\Delta u(t+j-d-1) + E_j(z^{-1})e(t+j) \quad (4-2)$$

این معادله را بصورت زیر بازنویسی می‌کنیم:

$$y(t+j) = F_j(z^{-1})y(t) + E_j(z^{-1})B(z^{-1})\Delta u(t+j-d-1) + E_j(z^{-1})e(t+j) \quad (5-2)$$

از آنجا که درجه معادله E_j برابر $j-1$ است، تمام جملات بیانگر نویز در معادله (5-2) مربوط به آینده هستند بنابراین بهترین تخمین برای $y(t+j)$ بصورت زیر بدست می‌آید:

$$\hat{y}(t+j|t) = G_j(z^{-1})\Delta u(t+j-d-1) + F_j(z^{-1})y(t) \quad (6-2)$$

که داریم:

$$G_j(z^{-1}) = E_j(z^{-1})B_j(z^{-1}) \quad (7-2)$$

جملات F_j و E_j بصورت زیر بیان می‌شوند:

$$F_j(z^{-1}) = f_{j0} + f_{j1}z^{-1} + \dots + f_{jn_a}z^{-n_a} \quad (8-2)$$

$$E_j(z^{-1}) = e_{j0} + e_{j1}z^{-1} + \dots + e_{j(j-1)}z^{-(j-1)} \quad (9-2)$$

نحوه استخراج F_{j+1} و E_{j+1} همانند F_{j+1} و E_{j+1} است یعنی ۱ را آنقدر بر $\tilde{A}(z^{-1})$ تقسیم می‌کنیم تا باقیمانده بصورت $z^{-(j+1)}F_{j+1}(z^{-1})$ فاکتورگیری شود:

$$F_{j+1}(z^{-1}) = f_{(j+1)0} + f_{(j+1)1}z^{-1} + \dots + f_{(j+1)n_a}z^{-n_a} \quad (10-2)$$

واضح است برای بدست آوردن F_{j+1} و E_{j+1} تنها لازمست عملیات استخراج F_j و E_j باندازه یک مرحله دیگر ادامه داده شوند. چندجمله‌ای E_{j+1} باینصورت خواهد شد:

$$E_{j+1}(z^{-1}) = E_j(z^{-1}) + e_{(j+1)j}z^{-j} \quad (11-2)$$

$$e_{(j+1)j} = f_{j0} \quad \text{که داریم:}$$

برای ضرایب چندجمله‌ای F_{j+1} داریم:

$$f_{(j+1)i} = f_{j(i+1)} - f_{j0}\tilde{a}_{i+1} \quad ; \quad i = 0, \dots, n_a - 1 \quad (12-2)$$

چندجمله‌ای G_{j+1} نیز بصورت بازگشتی محاسبه می‌شود:

$$G_{j+1} = E_{j+1}B = (E_j + f_{j0}z^{-j})B \quad (۱۳-۲)$$

$$G_{j+1} = G_j + f_{j0}z^{-j}B$$

بنابراین j ضریب اول چندجمله‌ای G_{j+1} مساوی ضرایب متناظر چندجمله‌ای G_j هستند و بقیه ضرایب از رابطه زیر بدست می‌آیند:

$$g_{(j+1)(j+i)} = g_{j(j+i)} + f_{j0}b_i \quad ; \quad i = 0, \dots, n_b \quad (۱۴-۲)$$

اکنون مجموعه پیش‌بینی‌های زیر را در نظر می‌گیریم:

$$\begin{aligned} \hat{y}(t+d+1|t) &= G_{d+1}\Delta u(t) + F_{d+1}y(t) \\ \hat{y}(t+d+2|t) &= G_{d+2}\Delta u(t+1) + F_{d+2}y(t) \\ &\vdots \\ \hat{y}(t+d+N|t) &= G_{d+N}\Delta u(t+N+1) + F_{d+N}y(t) \end{aligned} \quad (۱۵-۲)$$

این مجموعه را می‌توان بصورت زیر فشرده کرد:

$$y = GU + F(z^{-j})y(t) + G'(z^{-j})\Delta u(t-1) \quad (۱۶-۲)$$

که داریم:

$$y = \begin{bmatrix} \hat{y}(t+d+1|t) \\ \hat{y}(t+d+2|t) \\ \vdots \\ \hat{y}(t+d+N|t) \end{bmatrix} ; \quad U = \begin{bmatrix} \Delta u(t) \\ \Delta u(t+1) \\ \vdots \\ \Delta u(t+N) \end{bmatrix} ; \quad G = \begin{bmatrix} g_0 & 0 & \dots & 0 \\ g_1 & g_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ g_{N-1} & g_{N-2} & \dots & g_0 \end{bmatrix}$$

$$G'(z^{-1}) = \begin{bmatrix} (G_{d+1}(z^{-1}) - g_0)z \\ (G_{d+2}(z^{-1}) - g_0 - g_1z^{-1})z^2 \\ \vdots \\ (G_{d+N}(z^{-1}) - g_0 - g_1z^{-1} - \dots - g_{N-1}z^{-(N-1)})z^N \end{bmatrix} \quad (۱۷-۲)$$

دو جمله آخر در معادله (۱۶-۲) تنها به گذشته بستگی دارند بنابراین می‌توان هر دو را در یک جمله بنام f خلاصه کرد:

$$y = GU + f \quad (۱۸-۲)$$

در صورت صفر بودن همه شرایط اولیه پاسخ آزاد f نیز صفر خواهد بود. اگر پله واحد بصورت زیر به ورودی اعمال شود :

$$\Delta u(t) = 1, \Delta u(t+1) = 0, \dots, \Delta u(t+N-1) = 0$$

رشته خروجی مورد انتظار $[\hat{y}(t+1), \hat{y}(t+2), \dots, \hat{y}(t+N)]^T$ مساوی ستون اول ماتریس G می‌باشد. بنابراین ستون اول ماتریس G را می‌توان بصورت پاسخ فرایند به ازاء اعمال پله واحد به ورودی دستکاری‌شونده در نظر گرفت. پاسخ آزاد بصورت زیر محاسبه می‌شود :

$$f_{j+1} = z \left(1 - \bar{A}(z^{-1}) \right) f_j + B(z^{-1}) \Delta u(t-d+j) \quad (19-2)$$

که داریم :

$$\begin{aligned} f_0 &= y(t) \\ \Delta u(t+j) &= 0 \quad \text{for } j \geq 0 \end{aligned} \quad (20-2)$$

اکنون تابع هدف (۱-۱) را بصورت زیر می‌نویسیم :

$$J = (GU + f - w)^T (GU + f - W) + \lambda U^T U \quad (21-2)$$

که در این رابطه بردار W باین شکل تعریف می‌شود :

$$W = [w(t+d+1) \ w(t+d+2) \ \dots \ w(t+d+N)]^T \quad (22-2)$$

معادله (۲۱-۲) را بصورت زیر فشرده می‌کنیم :

$$J = \frac{1}{2} U^T H U + b^T U + f_0 \quad (23-2)$$

در رابطه بالا داریم :

$$\begin{aligned} H &= 2(G^T G + \lambda I) \\ b^T &= 2(f - W)^T G \\ f_0 &= (f - W)^T (f - W) \end{aligned} \quad (24-2)$$

با فرض عدم وجود قیود در سیگنال کنترل ، تابع هدف (۲۳-۲) با مساوی صفر قرار دادن گرادیان آن حداقل می‌شود و سیگنال حداقل‌کننده را بصورت زیر خواهیم داشت :

$$U = -H^{-1}b = (G^T G + \lambda I)^{-1} G^T (f - W) \quad (25-2)$$

طبق استراتژی افق پس‌رونده تنها عنصر اول از رشته کنترلی U به فرایند اعمال می‌شود:

$$\Delta u(t) = K(W - f) \quad (26-2)$$

در این رابطه K اولین سطر از ماتریس $(G^T G + \lambda I)^{-1} G^T$ است. حداقل افق پیش‌بینی برابر $d+1$ فرض شده و برای سهولت در علامت‌گذاریها، حداکثر افق پیش‌بینی را با N نمایش داده‌ایم. توجه شود بر خلاف کنترلرهای فیدبک کلاسیک که سیگنال کنترل با توجه به خطای گذشته محاسبه می‌شود در روش GPC تصمیم کنترلی بر اساس خطای پیش‌بینی شده برای آینده اتخاذ شده است.

۳-۲- محاسبه قانون کنترل GPC با در نظر گرفتن اغتشاشات قابل اندازه‌گیری

اکثر فرایندها متأثر از تغییرات خارجی ناشی از تغییر در متغیرهای قابل اندازه‌گیری می‌باشند. بعنوان مثال در راکتور پیوسته با ژاکت خنک‌کننده که درجه حرارت آن با تغییر دادن نرخ عبور جریان آب از ژاکت کنترل می‌شود هر گونه تغییر در نرخ عبور جریان باعث ایجاد تغییر در درجه حرارت راکتور می‌شود. این نوع اغتشاشات که اغتشاش بار نیز نامیده می‌شوند با استفاده از کنترلرهای پیش‌خور^۱ بسادگی قابل دستکاری می‌باشند. در MPC اغتشاشهای شناخته شده بصورت دقیق در نظر گرفته می‌شوند.

فرایندی را در نظر بگیرید که توسط مدل CARIMA با در نظر گرفتن اغتشاش توصیف شده است:

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + D(z^{-1})v(t) + \frac{C(z^{-1})\xi(t)}{\Delta} \quad (27-2)$$

که متغیر $v(t)$ اغتشاش اندازه‌گیری شده در زمان t است و $D(z^{-1})$ نیز بصورت زیر تعریف می‌شود:

$$D(z^{-1}) = d_0 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_{n_d} z^{-n_d} \quad (28-2)$$

با ضرب رابطه (۲۷-۲) در $\Delta E_j(z^{-1})z^j$ خواهیم داشت:

^۱ Feedforward

$$E_j(z^{-1})\tilde{A}(z^{-1})y(t+j) = E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + E_j(z^{-1})D(z^{-1})\Delta v(t+j) + E_j(z^{-1})e(t+j) \quad (29-2)$$

با انجام بعضی ساده‌سازیها به رابطه زیر می‌رسیم :

$$\hat{y}(t+j) = E(y(t+j)) = F_j(z^{-1})y(t) + E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + E_j(z^{-1})D(z^{-1})\Delta v(t+j) + E_j(z^{-1})e(t+j) \quad (30-2)$$

درجه $E_j(z^{-1})$ برابر $(j-1)$ است بنابراین جملات بیانگر نویز همگی مربوط به آینده خواهند بود. با استفاده از عملگر امید ریاضی و اطلاع از اینکه $E(e(t)) = 0$ ، مقادیر مورد انتظار برای $y(t+j)$ باین-صورت خواهند بود :

$$\hat{y}(t+j|t) = E(y(t+j)) = F_j(z^{-1})y(t) + E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + E_j(z^{-1})D(z^{-1})\Delta v(t+j) \quad (31-2)$$

اکنون چند جمله‌ای $H'_j(z^{-1})$ را مطابق رابطه زیر تعریف می‌کنیم :

$$E_j(z^{-1})D(z^{-1}) = H'_j(z^{-1}) + z^{-j}H'_j(z^{-1}) \quad (32-2)$$

با توجه به اینکه درجه چند جمله‌ای $H'_j(z^{-1})$ برابر $(j-1)$ می‌باشد، بعنوان معادله پیش‌بین خواهیم داشت :

$$\hat{y}(t+j|t) = G_j(z^{-1})\Delta u(t+j-1) + H'_j(z^{-1})\Delta v(t+j) + G'_j(z^{-1})\Delta u(t-1) + H'_j(z^{-1})\Delta v(t) + F_j(z^{-1})y(t) \quad (33-2)$$

توجه شود که سه جمله آخر سمت راست این معادله بستگی به مقادیر گذشته خروجی فرایند، اغتشاش قابل اندازه‌گیری و متغیرهای ورودی دارند که در صورت ثابت فرض کردن سیگنال کنترل و اغتشاشهای اندازه‌گیری شده متناظر با پاسخ آزاد فرایند خواهند بود در حالیکه جمله اول تنها به مقادیر آینده سیگنال کنترل بستگی داشته و می‌توان آنرا بعنوان پاسخ اجباری در نظر گرفت (پاسخی که با شرایط اولیه صفر بدست می‌آید : $y(t-j) = 0 ; \Delta u(t-j-1) = 0 ; \Delta v(t-j) = 0$ for $j > 0$)

جمله دوم عبارت (۲-۳۳) به اغتشاش قطعی^۱ بستگی دارد. در صورتیکه این اغتشاشات مربوط به بار فرایند باشد آنگاه مقادیر آینده آنها مشخص خواهد بود و در حالات دیگر باید مقادیر آینده آنها را پیش‌بینی کرد. اگر اغتشاشات بار در آینده ثابت و برابر آخرین مقدار اندازه‌گیری شده فرض شوند (یعنی $v(t+j) = v(t)$) جمله دوم حذف خواهد شد [۳۱].

معادله (۲-۲۲) را بصورت زیر بازنویسی می‌کنیم:

$$\hat{y}(t+j|t) = G_j(z^{-1})\Delta u(t+j-1) + H_j(z^{-1})\Delta v(t+j) + f_j \quad (۲-۳۴)$$

که داریم:

$$f_j = G'_j(z^{-1})\Delta u(t-1) + H'_j(z^{-1})\Delta v(t) + F_j(z^{-1})y(t) \quad (۲-۳۵)$$

اکنون مجموعه معادلات زیر را در نظر می‌گیریم:

$$\begin{aligned} \hat{y}(t+1|t) &= G_1 \Delta u(t) + H_1(z^{-1})\Delta v(t+1) + f_1 \\ \hat{y}(t+2|t) &= G_2 \Delta u(t+1) + H_2(z^{-1})\Delta v(t+2) + f_2 \\ &\vdots \\ \hat{y}(t+N|t) &= G_N \Delta u(t+N-1) + H_N(z^{-1})\Delta v(t+N) + f_N \end{aligned} \quad (۲-۳۶)$$

بخاطر خاصیت بازگشتی چندجمله‌ای E_j دستگاه معادلات (۲-۳۶) را می‌توان بصورت زیر نوشت:

$$\begin{bmatrix} \hat{y}(t+1|t) \\ \hat{y}(t+2|t) \\ \vdots \\ \hat{y}(t+j|t) \\ \vdots \\ \hat{y}(t+N|t) \end{bmatrix} = \begin{bmatrix} g_0 & 0 & \dots & 0 & \dots & 0 \\ g_1 & g_0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ g_{j-1} & g_{j-2} & \dots & g_0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{N-1} & g_{N-2} & \dots & \dots & \dots & g_0 \end{bmatrix} \begin{bmatrix} \Delta u(t) \\ \Delta u(t+1) \\ \vdots \\ \Delta u(t+j-1) \\ \vdots \\ \Delta u(t+N-1) \end{bmatrix} +$$

^۱ Deterministic

$$\begin{bmatrix} h_0 & 0 & \dots & 0 & \dots & 0 \\ h_1 & h_0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ h_{j-1} & h_{j-2} & \dots & h_0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & \dots & \dots & \dots & h_0 \end{bmatrix} \begin{bmatrix} \Delta u(t+1) \\ \Delta u(t+2) \\ \vdots \\ \Delta u(t+j-1) \\ \vdots \\ \Delta u(t+N) \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_j \\ \vdots \\ f_N \end{bmatrix} \quad (37-2)$$

که داریم $H_j(z^{-1}) = \sum_{i=1}^j h_i z^{-i}$. همچنین h_i ضرایب پاسخ سیستم به یک تغییر پله‌ای در اغتشاش

می‌باشند. با فرض $f' = HV + f$ ، قانون پیش‌بینی بصورت زیر درمی‌آید:

$$y = Gu + f' \quad (38-2)$$

فرم کلی معادله (۳۸-۲) همانند معادله پیش‌بین کلی مورد استفاده در حالت صفر بودن اغتشاشات اندازه‌گیری شده می‌باشد. پاسخ آزاد را بصورت پاسخ فرایند به شرایط اولیه (شامل اغتشاشات خارجی) و اغتشاشات معلوم آینده در نظر می‌گیریم و سیگنال کنترل آینده را با روش مشابهی محاسبه می‌کنیم.

۴-۲- محاسبه قانون کنترل بدون استفاده از معادله دیوفانتین

فرض می‌کنیم سیستم مورد نظر با استفاده از معادله زیر توصیف شود:

$$A(z^{-1})y(t) = B(z^{-1})z^{-d}u(t-1) \quad (39-2)$$

چند جمله‌ایهای $A(z^{-1})$ و $B(z^{-1})$ در معادله (۱۴-۱) تعریف شده‌اند. با استفاده از معادله (۳۹-۲) خواهیم داشت:

$$y(t) = (b_0 + \dots + b_{n_b} z^{-n_b})u(t-d-1) - (a_1 z^{-1} + \dots + a_{n_a} z^{-n_a})y(t) \quad (40-2)$$

$$y(t+1) = (b_0 z + b_1 + \dots + b_{n_b} z^{-n_b+1})u(t-d-1) - (a_1 + \dots + a_{n_a} z^{-n_a+1})y(t) \quad (41-2)$$

با کم کردن طرفین دو معادله (۴۰-۲) و (۴۱-۲) از یکدیگر به رابطه زیر می‌رسیم:

$$\begin{aligned}
 y(t+1) = & y(t) \\
 & + (b_0 z + (b_1 - b_0) + \dots + (b_{n_b} - b_{n_b-1}) z^{-n_b+1} + b_{n_b} z^{-n_b}) u(t-d-1) \\
 & + (-a_1 + (a_1 - a_2) z^{-1} + \dots + (a_{n_a+1} - a_{n_a}) z^{-n_a+1} + a_{n_a} z^{-n_a}) y(t)
 \end{aligned} \quad (42-2)$$

با استفاده از معادله بالا داریم :

$$\begin{aligned}
 \hat{y}(t+i) = & (1 - a_1) \hat{y}(t+i-1) + \sum_{j=1}^{n_a-1} (a_j - a_{j+1}) \hat{y}(t+i-1-j) \\
 & + \sum_{j=1}^{n_b} (b_j - b_{j+1}) u(t+i-1-d-j) + a_{n_a} \hat{y}(t+i-1-n_a) \\
 & + b_0 u(t+i-1-d)
 \end{aligned} \quad (43-2)$$

محاسبه پاسخ آزاد با استفاده از رابطه بالا ساده است.

۲-۵- انتخاب پارامترهای تابع هزینه

با توجه به تجربیات عملی بدست آمده در مورد پیاده‌سازی الگوریتم GPC، قواعد کلی زیر برای

انتخاب پارامترهای تابع هزینه بدست آمده‌اند (مرتبه مدل فرایند را با n نمایش داده‌ایم)

- $H_{1P} = n$: اگر زمان مرده سیستم مشخص باشد هیچ دلیلی وجود ندارد که حداقل افق پیش‌بینی را کمتر از آن انتخاب کنیم چون خروجی سیستم دارای زمان مرده در فاصله لحظه حاضر تا طی شدن زمان مرده تنها به ورودیهای قبل از این لحظه بستگی دارد و وارد کردن خروجی سیستم در بازه مذکور در محاسبات تنها حجم محاسبات را افزایش می‌دهد بدون اینکه اثر مثبتی روی دقت داشته باشد. پس H_{1P} باید بزرگتر از زمان مرده و حداقل مساوی مرتبه مدل فرایند باشد. اگر زمان مرده نامشخص یا متغیر باشد بدون از دست دادن پایداری، H_{1P} را برابر ۱ فرض کرده و درجه $B(z^{-1})$ (در مدل CARIMA) را آنقدر افزایش می‌دهیم تا از همه مقادیر ممکن زمان مرده بیشتر شود.
- $H_{2P} \geq (2n-1)$: اگر فرایند غیرمی‌نیمم فاز بوده و در ابتدا در جهت منفی حرکت کند H_{2P} باید بگونه‌ای انتخاب شود که نمونه‌های بعدی پاسخ فرایند که در جهت مثبت حرکت می‌کنند در تابع هزینه گنجانده شوند. این انتخاب در سیستمهای زمان‌گسسته بدان معنی است که H_{2P} از

درجه $B(z^{-1})$ بیشتر باشد اما در عمل مقادیر بزرگتری برای H_{2P} پیشنهاد می‌شود بطوریکه هر چه بیشتر به زمان برخاست^۱ سیستم نزدیکتر باشد.

- $H_C : (H_C \leq n) H_C$ یکی از پارامترهای مهم طراحی می‌باشد. برای یک سیستم ساده (مثلاً سیستم دارای پایداری حلقه باز، حتی اگر زمان مرده داشته باشد یا اینکه غیرمی‌نیمم فاز باشد)، $H_C = 1$ کنترل قابل قبولی ارائه می‌کند. افزایش H_C از یکطرف باعث فعالتر شدن سیگنالهای کنترل و پاسخ خروجی متناظر می‌شود و از طرف دیگر حجم محاسبات را بالا می‌برد. برای سیستمهای پیچیده‌تر، H_C باید بزرگتر بوده و حداقل مساوی تعداد قطبهای ناپایدار یا زیرمیرا باشد. در حالت کلی مقادیر کوچک برای H_C باعث کاهش قابل توجه هزینه محاسبات بهینه‌سازی می‌شود.

۲-۶- فرمول‌بندی روش GPC در فضای حالت

بعلت سهولت طراحی کنترلر برای فرم فضای حالت، این بخش به فرمول‌بندی روش GPC در فضای حالت اختصاص یافته است.

۲-۶-۱- مدل فضای حالت بدون اغتشاش

فرض کنیم که سیستم تحت کنترل با معادله زیر بیان شده باشد:

$$x(t+1) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t)$$

معادله پیش‌بین حالت با استفاده از مدل فوق بصورت زیر درمی‌آید:

$$\hat{x}(t+j|t) = A\hat{x}(t+j-1|t) + Bu(t+j-1|t)$$

همچنین روابط زیر را بعنوان معادلات پیش‌بین خروجی خواهیم داشت:

$$\hat{y}(t+1|t) = CAx(t) + CBu(t)$$

$$\hat{y}(t+2|t) = CA^2x(t) + CABu(t) + CBu(t+1)$$

⋮

$$\hat{y}(t+H_{2P}|t) = CA^{H_{2P}}x(t) + CA^{H_{2P}-1}Bu(t) + \dots + CA^{H_{2P}-H_C-1}Bu(t+H_C)$$

^۱. Rise Time

معادلات بالا باین فرم فشرده می‌شوند :

$$Y = Gx(t) + F_1 U \quad (44-2)$$

$$Y = [\hat{y}(t+1|t) \hat{y}(t+2|t) \dots \hat{y}(t+H_{2P}|t)]^T \quad (45-2)$$

$$G = [CA \ CA^2 \ \dots \ CA^{H_{2P}}]^T \quad (46-2)$$

$$F_1 = \begin{bmatrix} CB & 0 & \dots & 0 \\ CA^2 B & CB & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ CA^{H_{2P}-1} B & CA^{H_{2P}-2} B & \dots & CA^{H_{2P}-H_C-1} B \end{bmatrix} \quad (47-2)$$

$$U = [u(t) \ u(t+1) \ \dots \ u(t+1)]^T \quad (48-2)$$

تابع هدف را بصورت زیر در نظر می‌گیریم :

$$J = \frac{1}{2} \sum_{j=1}^{H_{2P}} [y_{ref}(t+j) - \hat{y}(t+j)]^T Q_j [y_{ref}(t+j) - \hat{y}(t+j)] + \frac{1}{2} \sum_{j=1}^{H_C} u(t+j)^T R_j u(t+j) \quad (49-2)$$

بفرم برداری داریم :

$$J = \frac{1}{2} (Y_{ref} - Y)^T Q (Y_{ref} - Y) + U^T R U \quad (50-2)$$

که $Q \geq 0$ بوده و R یک ماتریس متقارن می‌باشد (ماتریسهای وزنی). با جایگزینی (44-2) در (50-2) و حداقل‌سازی J رسته کنترل بصورت زیر بدست می‌آید :

$$U = (F_1^T Q F_1 + R)^{-1} F_1^T Q [Y_{ref} - Gx(t)] \quad (51-2)$$

تنها عنصر اول از رسته کنترلی محاسبه شده به سیستم اعمال می‌شود بنابراین خواهیم داشت :

$$u(t) = N_1 [Y_{ref} - Gx(t)] \quad (52-2)$$

که $Q = F_1^T (F_1^T Q F_1 + R)^{-1} F_1^T$ در حقیقت n سطر اول ماتریس $N_1 = [I_n, 0, 0, \dots, 0]$ می‌باشد (n بعد ورودی است). می‌توانیم تابع هدف را بصورت زیر نیز تعریف کنیم:

$$J = \frac{1}{2} \sum_{j=1}^{H_{2P}} \left[y_{ref}(t+j) - \hat{y}(t+j) \right]^T Q_j \left[y_{ref}(t+j) - \hat{y}(t+j) \right] + \frac{1}{2} \sum_{j=1}^{H_C} \Delta u(t+j)^T R_j \Delta u(t+j) \quad (53-2)$$

که دارای فرم برداری زیر خواهد بود.

$$J = \frac{1}{2} (Y_{ref} - Y)^T Q (Y_{ref} - Y) + \Delta U^T R \Delta U \quad (54-2)$$

در این حالت ابتدا معادله (۴۴-۲) را بفرم زیر می‌نویسیم:

$$Y = Gx(t) + F_{11} \Delta U + F_2 u(t-1) \quad (55-2)$$

که در این معادله داریم:

$$F_{11} = \begin{bmatrix} CB & 0 & \dots & 0 \\ C(A+I)B & CB & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ C\gamma(H_{2P})B & C\gamma(H_{2P}-1)B & \dots & CA^{H_{2P}-H_C-1}B \end{bmatrix} \quad (56-2)$$

$$\Delta U = [\Delta u(t) \ \Delta u(t+1) \ \dots \ \Delta u(t+H_{2P})]^T \quad (57-2)$$

$$F_2 = \begin{bmatrix} CB \\ C(A+I)B \\ \vdots \\ C(A^{H_{2P}-1} + A^{H_{2P}-2} + \dots + A^{H_{2P}-H_C-1})B \end{bmatrix} \quad (58-2)$$

$$\gamma(j) = \sum_{k=0}^{k=j} A^{(H_{2P}-H_C-1+k)} \quad (59-2)$$

با جایگزینی (۵۵-۲) در (۵۴-۲) و حداقل‌سازی تابع هدف داریم:

$$\Delta U = (F_{11}^T Q F_{11} + rI)^{-1} F_{11}^T Q [Y_d - Gx(t) - F_2 u(t-1)] \quad (۶۰-۲)$$

بنابراین بعنوان سیگنال کنترل اعمالی به سیستم خواهیم داشت :

$$\Delta u(t) = N_1 [Y_d - Gx(t) - F_2 u(t-1)] \quad (۶۱-۲)$$

که $N_1 = [I_m, 0, 0, \dots, 0]$ بیانگر n سطر اول ماتریس $(F_{11}^T Q F_{11} + R)^{-1} F_{11}^T Q$ است.

۲-۶-۲- مدل فضای حالت با در نظر گرفتن اغتشاش

وجود اغتشاش در سیستم باعث ایجاد پیچیدگی و محدودیت در عملکرد الگوریتم GPC بمیزان قابل توجهی می‌شود. برای گروه بزرگی از اغتشاشهای متناوب و تکرارشونده می‌توانیم الگوریتم GPC را با استفاده از یک طرح یادگیری مناسب بگونه‌ای بهبود بخشیم که اثرات نامطلوب اینگونه اغتشاشها تا حد امکان حذف شوند. با این انگیزه در ادامه به معرفی فرمول‌بندی جدیدی می‌پردازیم که در آن روش GPC با یک طرح پیش‌خور یادگیری اغتشاش^۱ که مبتنی بر ایده کنترل یادگیری تکرارشونده^۲ می‌باشد ترکیب شده است. ایده اولیه طرح کنترل اینست که با فیدبک کردن خطاهای کنترل قبلی، سیگنال کنترل را برای عملکرد حال حاضر بهبود بخشیم.

فرض کنیم که سیستم تحت کنترل با معادله زیر بیان شده باشد :

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + v(t) \end{aligned} \quad (۶۲-۲)$$

در این مدل $w(t)$ اغتشاش ورودی است که تکرارپذیر فرض می‌شود و $v(t)$ نامعینی خروجی می‌باشد. معادله پیش‌بین حالت با استفاده از مدل فوق بصورت زیر درمی‌آید:

$$\hat{x}(t+j|t) = A\hat{x}(t+j-1|t) + Bu(t+j-1|t) \quad (۶۳-۲)$$

و برای معادلات پیش‌بین خروجی خواهیم داشت :

^۱ Feedforward Disturbance Learning Scheme
^۲ Iterative Learning Control

$$\hat{y}(t+1|t) = CAx(t) + CBu(t) + Cw(t)$$

$$\hat{y}(t+2|t) = CA^2x(t) + CABu(t) + CBu(t+1) + CAw(t) + Cw(t+1)$$

$$\vdots$$

$$\hat{y}(t+H_{2P}|t) = CA^{H_{2P}}x(t) + CA^{H_{2P}-1}Bu(t) + \dots + CA^{H_{2P}-H_C-1}Bu(t+H_C) \\ + CA^{H_{2P}-1}w(t) + \dots + Cw(t+H_{2P}-1)$$

با فرض ثابت بودن اغتشاش در طول مدت افق پیش‌بینی معادلات بالا بفرم زیر فشرده می‌شوند:

$$Y = Gx(t) + F_{11}\Delta U + F_2u(t-1) + F_3w(t) \quad (64-2)$$

F_3 بصورت زیر تعریف می‌شود:

$$F_3 = \begin{bmatrix} C \\ C(A+I) \\ \vdots \\ C(A^{H_{2P}-1} + A^{H_{2P}-2} + \dots + I) \end{bmatrix} \quad (65-2)$$

تابع هدف (۵۳-۲) را با فرض $H_C < H_{2P}-1$ در نظر می‌گیریم. رشته کنترل حداقل‌کننده این تابع هزینه بصورت زیر بدست می‌آید:

$$\Delta U = (F_1^T Q F_1 + R)^{-1} F_1^T Q [Y_{ref} - Gx(t) - F_2w(t) - F_3u(t-1)] \quad (66-2)$$

که در این رابطه داریم:

$$Y_{ref} = [y_{ref}(t+1) \ y_{ref}(t+2) \ \dots \ y_{ref}(t+H_{2P})]^T \quad (67-2)$$

سیگنال کنترل اعمالی به سیستم بصورت زیر بدست می‌آید:

$$\Delta u(t) = N_1 [Y_{ref} - Gx(t) - F_2w(t) - F_3u(t-1)] \quad (68-2)$$

۲-۶-۳- مدل فضای حالت با رویتگر

نیاز به اندازه‌گیری حالت‌های سیستم یکی از محدودیتهای اصلی الگوریتم GPC مبتنی بر فضای حالت می‌باشد. در این بخش یک کنترلر- رویتگر پیش‌بین تعمیم‌یافته برای سیستم‌های دارای حالات غیرقابل اندازه‌گیری و تاخیردار ارائه می‌شود که از فیلتر کالمن بعنوان رویتگر حالت استفاده می‌کند.

فرض شود که سیستم مورد نظر با معادلات زیر توصیف شده است.

$$x(t+1) = Ax(t) + Bu(t-d) + D\xi(t) \quad (۶۹-۲)$$

$$y(t) = Cx(t) + \xi(t) \quad (۷۰-۲)$$

در این رابطه d تاخیر زمانی ورودی، ξ نویز سفید با میانگین صفر و ماتریس کواریانس Σ است. بعنوان مدل پیش‌بین حالت داریم:

$$\hat{x}(t+j|t) = A\hat{x}(t+j-1|t) + Bu(t+j-1|t) \quad ; \quad j=1, \dots, H_C+d+1 \quad (۷۱-۲)$$

تابع هدف نیز بصورت زیر در نظر گرفته می‌شود:

$$J = \frac{1}{2} \sum_{j=1}^{H_C+d+1} \left[y_{ref}(t+j) - \hat{y}(t+j) \right]^T Q_j \left[y_{ref}(t+j) - \hat{y}(t+j) \right] + \frac{1}{2} \sum_{j=1}^{H_C} \Delta u(t+j)^T R_j \Delta u(t+j) \quad (۷۲-۲)$$

با جایگذاری مکرر فرم بازگشتی معادله (۶۹-۲) در (۷۰-۲) خواهیم داشت:

$$\hat{y}(t+1|t) = C\hat{x}(t+1|t)$$

$$\hat{y}(t+2|t) = CA\hat{x}(t+1|t) + CBu(t-d+1)$$

⋮

$$\hat{y}(t+d+H_C+1|t) = CA^{H_C+d}\hat{x}(t+1|t)$$

$$+ CA^{H_C+d-1}Bu(t-d+1) + \dots + CBu(t+H_C)$$

فرم فشرده معادلات بالا بدین شکل می‌باشد:

$$\hat{Y} = G\hat{x}(t+1|t) + F_1\bar{U}_t + \bar{F}_2\Delta u(t) \quad (۷۳-۲)$$

که داریم:

$$G = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{H_C+d} \end{bmatrix}; F_1 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ CB & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ CA^{d-1}B & CB & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ CA^{H_C+d-1}B & CA^{H_C+d-2}B & \dots & CA^{d+1}B \end{bmatrix}$$

$$\bar{U}_t = [u(t) u(t+1) \dots u(t+H_C)]^T$$

$$\bar{F}_2 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ CB \\ \vdots \\ CA^{H_C}B + CA^{H_C-1}B + \dots + CB \end{bmatrix}$$

اکنون فیلتر کالمن را با معادلات زیر در نظر می‌گیریم:

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t-d) + K(t)(y(t) - H\hat{x}(t)) \quad (74-2)$$

$$\hat{x}(t_0) = E\{x(t_0)\}$$

و فرض می‌کنیم که

$$e(t) = x(t) - \hat{x}(t) \quad (75-2)$$

$$p(t) = E\{e(t)e^T(t)\}$$

با تفریق (74-2) از (75-2) مشخص می‌شود که خطای تخمین حالت در رابطه زیر صدق می‌کند:

$$e(t+1) = (A - K(t)H)e(t) + (D - K(t))\xi(t) \quad (76-2)$$

ماتریس بهره K بصورت زیر بدست می‌آید:

$$K(t) = (AP(t)C^T + D\xi) (CP(t)C^T + \xi)^{-1} \quad (77-2)$$

همچنین رابطه بازگشتی زیر را برای $p(t)$ خواهیم داشت:

$$p(t+1) = (A - K(t)H)p(t)A^T + D D^T \xi - K(t)D^T \xi \quad (78-2)$$

اکنون انتخابهای زیر را در نظر می‌گیریم :

$$R_0=r \quad ; \quad R_j=0 \quad ; \quad j=1,\dots,H_C+d \quad (79-2)$$

با استفاده از این انتخابها تابع هدف (72-2) بصورت زیر درمی‌آید :

$$J = \frac{1}{2} \sum_{j=1}^{H_C+d+1} \left[y_{ref}(t+j) - \hat{y}(t+j) \right]^T Q_j \left[y_{ref}(t+j) - \hat{y}(t+j) \right] + \frac{1}{2} \Delta u(t)^2 \quad (80-2)$$

این تابع هدف توسط ورودی زیر حداقل می‌شود :

$$\Delta u(t) = \left(\bar{F}_2^T \bar{F}_2 + r \right)^{-1} \bar{F}_2^T \left[Y_{ref} - G(A - KH) \hat{x}(t) - \bar{F}_1 \bar{U}_t \right] \quad (81-2)$$

که در این رابطه داریم :

$$Y_{ref} = \left[y_{ref}(t+1) \quad y_{ref}(t+1) \quad \dots \quad y_{ref}(t+H_C+1) \right]^T \quad (82-2)$$

$$K = APC^T \left(CPC^T + \zeta \right)^{-1} \quad (83-2)$$

$$\bar{F}_1 = \left[GB \quad F_1 + \bar{F}_2 \times \overbrace{[0 \ 0 \ \dots \ 1]}^{h-1} \right] \quad (84-2)$$

$$\bar{U}_t = \left[u(t-d) \quad u(t-d+1) \quad \dots \quad u(t-1) \right]^T \quad (85-2)$$

فصل سوم

شناسائی برخط

مدل فازی TS

۳-۱- مقدمه

گفته شد که روشهای MPC مدل سیستم را بطور مستقیم در طرح کنترلی بکار می‌گیرند بنابراین مدل نقش مستقیمی در طراحی و کیفیت عملکرد کنترلر خواهد داشت. صرفنظر از طرح کنترلی یکی از پیش‌نیازهای اصلی برای طراحی کنترلر داشتن یک مدل مناسب از سیستم می‌باشد. بدلیل وجود بعضی نواقص و مشکلات در مدل‌های متداولی که بطور مختصر در فصل اول بیان شدند، فصل سوم به شناسایی برخط مدل فازی تاکاگی- سوگنو (TS) اختصاص یافته است.

۳-۲- بررسی مختصر روشهای شناسایی سیستم فازی و مشکلات آنها

در مراجع [۱۲-۱۳] ثابت شده است که سیستمهای فازی تقریب‌زنده‌های عمومی هستند این گرایش عمدتاً بدلیل معنی‌دار و مستدل بودن منطق فازی است. همچنین فهم قواعد فازی و کارکردن با آنها نسبتاً ساده است و امکان نمایش صریح دانش زبانی کاربران بفرم قواعد اگر-آنگاه وجود دارد. از بین سیستمهای فازی مختلف، کلاس TS به ابزار قدرتمندی در مدلسازی و کنترل سیستمهای پیچیده و دارای دینامیک متغیر تبدیل شده است. در حالیکه روش جدول‌بندی بهره از یک مجموعه مدل خطی برای تقریب محلی سیستم غیرخطی استفاده می‌کند و خطی‌سازی را دقیقاً در نقاط کار مشخص انجام می‌دهد در مدل‌های TS، ایده خطی‌سازی محلی در نواحی تعریف شده بصورت فازی بکار گرفته می‌شود. با توجه به فازی بودن نواحی، سیستم غیرخطی بصورت چند زیرمدل خطی مدلسازی می‌شود که هر کدام از این زیرمدلها به یکی از نواحی کاری سیستم تعلق دارد و خروجی نهائی مدل با توجه به خروجی تک‌تک زیرمدلها محاسبه می‌شود.

در سیستمهای فازی-عصبی از الگوریتمهای یادگیری شبکه‌های عصبی برای استخراج مدل‌های فازی از داده‌ها استفاده می‌شود. این سیستمها خصوصیات یادگیری شبکه‌های عصبی و مزایای منطق فازی را با هم جمع کرده‌اند و جزئیات مربوط به آنها را می‌توان در مرجع [۲۶] یافت.

دو مرحله اساسی در روشهای فازی-عصبی، شناسایی ساختار مدل و تعیین پارامترها (بهینه‌سازی) می‌باشند. در مرحله شناسایی ساختار، تعداد قواعد و نقاط کانونی آنها (مراکز و عرضهای

توابع عضویت مجموعه‌های فازی مقدمه) توسط خوشه‌سازی فازی تخمین زده می‌شوند. با ثابت ماندن پارامترهای مقدمه، مدل TS بصورت خطی درمی‌آید که پارامترهای قسمت نتیجه در مرحله شناسائی پارامتر بدست می‌آیند. در اکثر روشهای متداول فازی-عصبی از الگوریتم پس‌انتشار BP (اجرا شده بوسیله روش کاهش گرادیان^۱) یا روشهای مشابه آن برای بهینه‌سازی پارامترهای قسمت نتیجه قواعد استفاده می‌شود.

در [۲۲] یک روش سلسله مراتبی بهبود یافته جهت کاربرد در سیستمهای فازی-تطبیقی ارائه شده است که مبتنی بر الگوریتم یادگیری خودسازمانده می‌باشد. متاسفانه الگوریتم بصورت خارج خط بوده و همه پارامترهای آن بر اساس الگوریتم مبتنی بر کاهش گرادیان آموزش می‌بینند که این امر افزایش حجم محاسبات و کاهش سرعت همگرایی را بدنبال دارد. در [۲۳] شبکه‌های شعاعی بازگشتی جهت حذف نویز تطبیقی بکار رفته‌اند و برای شناسائی ساختار (اختصاص مراکز) از الگوریتم خوشه-سازی میانگین K استفاده شده است که تنها جهت یادگیری خارج خط مناسب می‌باشد .

الگوریتم ترتیبی ارائه شده در [۲۴] قابلیت کاربرد در سیستمهای بلادرنگ را دارا می‌باشد اما نمونه‌های قبلی داده‌ها باید در یک پنجره لغزان جهت ایجاد نوروهای پنهان ذخیره شوند که این مورد با اصول و نیازهای پردازش موازی در تناقض است. در مراجع [۱۹] و [۲۵] شبکه‌های فازی-عصبی دینامیک با ساختار تنظیم‌شونده جهت کاربرد در تقریب توابع، شناسائی سیستم و پیش‌بینی معرفی شده‌اند. جهت تعیین پارامترهای قسمت نتیجه و تنظیم دینامیک ساختار باید تمام داده‌های آموزشی قبلی ذخیره شوند در نتیجه حجم محاسبات و حافظه مورد نیاز بطرز نامطلوبی افزایش می‌یابد .

در مرجع [۱۴] روشی برای استخراج مدل‌های فازی-عصبی تطبیقی ارائه شده است که در آن شناسائی ساختار و شناسائی پارامتر با یکدیگر ترکیب شده‌اند. در مرجع [۱۵] روشهایی برای استخراج مدل‌های فازی TS دقیق از خوشه‌های فازی ایجاد شده بوسیله الگوریتم خوشه‌سازی Gath-Geva معرفی شده‌اند و در مراجع [۱۶] و [۱۷] از الگوریتم ژنتیک برای آموزش ساختار مدل و پارامترهای آن استفاده شده است. مشکل مشترک همگی روشهای ذکر شده تاکنون اینست که کل داده‌ها باید در

^۱ GD: Gradient Decent

ابتدای آموزش در دسترس باشند همچنین بازآموزی خارج خط این روشها تکراری و وقت گیر بوده و کاربرد آنها را در سیستمهای بلادرنگ بسیار مشکل یا غیرممکن می سازد. روشهای مذکور تنها در مسائل قطعی^۱ می توانند عملکرد مناسبی داشته باشند و در مورد مسائل استوکاستیک کارکرد خوبی نخواهند داشت .

طی دو دهه گذشته با پیشرفت سریع تکنولوژی اطلاعات ، حجم اطلاعات نیز بشدت افزایش یافته است. این اطلاعات حجیم و غیرایستا باعث بوجود آمدن چالشهایی در روشهای یادگیری آماری شده اند [۳] تحلیل جریانهای داده ها^۲ با روشهای خارج خط (دسته ای^۳) امکان پذیر نیست چون نمی توان همه این داده ها را ذخیره کرد . بنابراین نیاز به روشهایی داریم که بجای ذخیره کردن خود داده ها ، اطلاعات موجود در داده ها را بصورت یک ساختار درک شدنی و تعبیر پذیر^۴ در خود ذخیره کنند . با توجه به غیرایستا بودن داده ها ساختار مدل باید دینامیک و تکاملی^۵ باشد . منظور از مدل فازی تکاملی مدلی است که ساختار (پایگاه قواعد) و پارامترهای آن بصورت تدریجی توسعه و تکامل پیدا کنند. در [۴] روش جدیدی برای یادگیری برخط مدل های TS معرفی شده است که ساختار مدل TS و پارامترهای آنرا بصورت بازگشتی بروز می کند. این روش جهت کاربرد در مدلسازی غیرخطی و کنترل تطبیقی بسیار جذاب بنظر می رسد اما در بعضی حالات ممکن است تعداد قواعد فازی خیلی زیاد شوند چون معیاری برای حذف قواعد اضافی از پایگاه قواعد در نظر گرفته نشده است . همچنین شعاع تاثیر قواعد فازی (شعاع خوشه ها) ثابت در نظر گرفته شده است و باید در ابتدای آموزش انتخاب شود که این انتخاب با توجه به اهمیت زیاد شعاع خوشه ها و ثابت نبودن شعاع واقعی انتخاب مطلوبی بنظر نمی رسد . همین مشکل در [۱] نیز دیده می شود .

در این فصل به معرفی یک روش شناسائی برخط تطبیقی برای استخراج مدل TS تکاملی سیستم ناشناخته MIMO می پردازیم که در آن شناسائی ساختار و پارامترهای مدل TS بصورت برخط انجام می شود . این مدل که مبتنی بر پایگاه قواعد تکاملی^۶ (ER) است از پتانسیل اطلاعاتی^۱

^۱ Deterministic

^۲ Data Stream

^۳ Batch Mode

^۴ Understandable And Interpretable

^۵ Evolving

^۶ Evolving Rule-Based Models

داده جدید بعنوان تریگری برای بروزرسانی قواعد استفاده می‌کند. پتانسیل اطلاعاتی هر نمونه نمایش‌دهنده میزان مجاورت فضائی^۲ داده‌ها با آن نمونه (فاصله نمونه مذکور از تمام نمونه‌های قبلی) بوده و برای تعریف نقاط کانونی قواعد (پارامترهای مقدمه) بکار می‌رود. علت استفاده از پتانسیل بجای فاصله اینست که بر خلاف فاصله، پتانسیل اطلاعاتی می‌تواند اطلاعات فاصله فضائی تجمعی مربوط به تمام داده‌های قبلی را در خود ذخیره کند. جهت اطمینان از فشرده‌بودن پایگاه قواعد، معیارهای ابتکاری برای تشخیص و حذف قواعد اضافی گنجانده شده است و شعاع تاثیر قواعد (خوشه‌ها) در طی آموزش بوسیله الگوریتم کاهش گرادیان اصلاح می‌شود تا از کافی بودن همپوشانی بین خوشه‌ها اطمینان حاصل شود. همچنین شرایط لازم و کافی برای پایدار بودن الگوریتم کاهش گرادیان نیز ارائه شده‌اند.

۳-۳- معرفي مدل فازی TS و شبکه فازی-عصبی

۳-۳-۱- معرفي مدل فازی TS

در مدل‌های TS بخش‌های مقدمه قواعد بصورت فازی و بخش‌های نتیجه قواعد، توابعی خطی از ورودیها هستند و از روش استدلال تاکاگی-سوگنو-کانگ تبعیت می‌کنند. در یک مدل TS دارای n ورودی و یک خروجی، قاعده i ام بصورت زیر بیان می‌شود:

$$R_i : IF(x_1 \text{ is } A_{i1}) AND \dots AND(x_n \text{ is } A_{in}) \quad (1-3)$$

$$THEN y^i = a_0^i + a_1^i x_1 + \dots + a_n^i x_n = x_e^T \pi^i \quad ; \quad i = \{1, R\}$$

A_{ij} : مجموعه‌های فازی قسمتهای مقدمه ($j = \{1, n\}$)

x : بردار ورودی $x = [x_1, x_2, \dots, x_n]^T$

y_i : خروجی i امین زیرمدل خطی

a_{il} : پارامترهای بخش نتیجه زیرمدل خطی i ام ($l = \{0, n\}$)

همچنین R تعداد قواعد فازی و $x_e = [1, x^T]^T$ بردار ورودی گسترش‌یافته می‌باشد. برای

سیستم چند ورودی-یک خروجی داریم:

^۱ Informative Potential

^۲ Spatial Proximity Measure

$$\pi^i = [a_{00}^i, a_{01}^i, \dots, a_{0n}^i]^T \quad (2-3)$$

در مدل‌سازی سیستم چندورودی-چندخروجی بردار y^i بصورت چندبعدی درمی‌آید که با فرض m متغیر خروجی خواهیم داشت :

$$y^i = [y_1^i, y_2^i, \dots, y_m^i] \quad (3-3)$$

بردار پارامترهای بخش مقدمه نیز باینصورت بیان می‌شود :

$$\pi^i = \begin{bmatrix} a_{01}^i & a_{02}^i & \dots & a_{0m}^i \\ a_{11}^i & a_{12}^i & \dots & a_{1m}^i \\ \dots & \dots & \dots & \dots \\ a_{n1}^i & a_{n2}^i & \dots & a_{nm}^i \end{bmatrix} \quad (4-3)$$

گفته شد که در مدل TS بجای خطی‌سازی در نقاط کار مشخص از ایده خطی‌سازی در نواحی تعریف شده بصورت فازی استفاده می‌شود یعنی فضای داده‌ها به تعدادی ناحیه تقسیم‌بندی شده و برای تعریف هر ناحیه بصورت فازی برای آن پارامترهایی در نظر می‌گیریم که هر ناحیه با یک زیرمدل خطی متناظر می‌شود. با توجه به اینکه مقدمه‌های قواعد بصورت فازی بیان شده‌اند سیستم غیرخطی ، بصورت مجموعه‌ای از زیرمدلهای خطی مدل‌سازی می‌شود که بصورت فازی (آزادانه) به یکدیگر کوپل شده‌اند. سطح فعال‌سازی هر قاعده ، سهم زیرمدل خطی متناظر با این قاعده در تعیین خروجی نهائی مدل TS را مشخص می‌کند.

بعلت اینکه ارتباط خروجی نهائی مدل و خروجیهای زیرمدلها بصورت فازی است مدل کلی غیرخطی خواهد بود. قدرت مدل‌های غیرخطی در نمایش رفتار سیستمها بیشتر از مدل‌های خطی است اما از طرف دیگر شناسایی مدل‌های غیرخطی و طراحی کنترلر برای آنها دشوارتر از مدل‌های خطی می‌باشد. اما مدل غیرخطی TS در عین داشتن دقت بالا ، بعلت خطی بودن زیرمدلهای تشکیل‌دهنده آن تا حد قابل توجهی از این قاعده مستثنی می‌باشد (بخصوص در مورد طراحی کنترلر) . در کنترلرهای پیش‌بین یک مساله بهینه‌سازی در هر پله زمانی باید حل شود. در صورت کاربرد مدل‌های غیرخطی مساله مذکور نیز غیرخطی می‌شود که جواب تحلیلی برای آن وجود ندارد و باید با استفاده از روشهای وقت‌گیر مبتنی بر تکرار حل شود. اما در مورد مدل غیرخطی TS خطی بودن زیرمدلها باعث می‌شود که بجای یک مساله بهینه‌سازی غیرخطی چندین مساله بهینه‌سازی خطی داشته باشیم که جوابهای

تحلیلی برای آنها موجود می‌باشد در نتیجه حجم محاسبات کاهش یافته و پیاده‌سازی سخت‌افزاری کنترلر آسانتر می‌شود.

یافتن مدل فازی در حقیقت معادل با یافتن ارتباط بین عناصر بردار ورودی و نمونه بعدی خروجی است بنابراین در انتخاب عناصر بردار ورودی باید از وجود این ارتباط منطقی مطمئن شویم در غیر اینصورت قرار دادن نمونه‌های اضافی در بردار ورودی تنها حجم محاسبات را افزایش داده و اثر بسیار ناچیزی روی دقت مدل خواهد داشت. بعنوان مثال در مورد سیستمهای تاخیردار قرار دادن نمونه‌های جدید ورودی در بردار ورودی مفید نیست. یک بردار ورودی نوعی برای سیستمهای تاخیردار بصورت زیر می‌باشد:

$$z = [y(t) \ y(t-1) \dots \ y(t-m), u(t-d) \ u(t-d-1) \dots \ u(t-d-n-1)] \quad (۵-۳)$$

سیستم فازی با استفاده از یک شبکه عصبی پیش‌خور^۱ اجرا می‌شود.

۳-۳-۲- ساختار شبکه فازی-عصبی

- لایه اول (ورودی): تعداد گره‌های این لایه مساوی تعداد متغیرهای ورودی (n) است و هیچ محاسبه‌ای در آنها انجام نمی‌شود بلکه تنها مقادیر ورودیها را به لایه دوم انتقال می‌دهند.
- لایه دوم (توابع عضویت): گره‌های این لایه به R گروه تقسیم می‌شوند که هر گروه شامل n گره است و هر کدام از گره‌ها تابع عضویت یکی از مجموعه‌های فازی بخشهای مقدمه قواعد را محاسبه می‌کنند. توابع عضویت می‌توانند از نوع مثلثی، ذوزنقه‌ای، گوسی و غیره باشند که در اینجا بدون از دست دادن کلیت مطلب از توابع عضویت گوسی بصورت زیر استفاده می‌کنیم:

$$\mu_{ij}(x_j) = \exp\left(-\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}\right); \quad i = \{1, R\} \quad j = \{1, n\} \quad (۶-۳)$$

c_{ij} : مرکز تابع عضویت گوسی مجموعه فازی A_{ij} (زامین عنصر مرکز خوشه i ام)

σ_{ij} : عرض تابع عضویت گوسی مجموعه فازی A_{ij} (زامین عنصر شعاع خوشه i ام)

^۱. Feedforward

علت انتخاب تابع گوسی اینست که اولاً این تابع هموار و نرمال بوده و همه جا مشتق پذیر است که خاصیت مشتق پذیری برای سیستمهای فازی-عصبی که از روشهای یادگیری مبتنی بر کاهش گرادیان استفاده می کنند مفید می باشد. ثانیاً تابع گوسی قادر است بیشترین عمومیت^۱ ممکن را در توصیف مجموعه های فازی ایجاد کند (تابع گوسی توزیع داده ها در خوشه را بطور نرمال در نظر می گیرد). همچنین تابع گوسی چندبعدی را براحتی می توان بصورت ضرب چندین تابع عضویت یک-بعدی تجزیه کرد .

بعضی مراجع از مقادیر یکسانی برای پهنای توابع عضویت متغیرهای مختلف در هر قاعده استفاده کرده اند یعنی در رابطه (۴-۶) بجای σ_{ij} از σ_i استفاده شده است . از نقطه نظر هندسی این بدان معناست که بعنوان مثال در فضای دوبعدی ، خوشه ها بصورت دایره ای و در فضای سه بعدی بصورت کروی در نظر گرفته شده اند. در حالیکه توزیع واقعی داده ها ممکن است شبیه به بیضی ، مستطیل یا هر شکل هندسی دیگر باشد بنابراین بهتر است عرض تابع عضویت هر متغیر خوشه بصورت جداگانه در نظر گرفته شود .

در [۴] تابع عضویت بصورت زیر تعریف شده است :

$$\mu_{ij}(x_j) = \exp\left(-4 \frac{(x_j - c_{ij})^2}{r_j^2}\right) ; i = \{1, R\} \quad j = \{1, n\} \quad (7-3)$$

r_j ثابت مثبتی است که ناحیه تاثیر مدل (شعاع همسایگی نقطه کانونی قاعده) را مشخص می کند . در اینجا عرض توابع عضویت هر متغیر در قواعد مختلف بصورت یکسان فرض شده است . شعاع r_j یکی از محدود پارامترهایی است که باید قبل از شروع الگوریتم مشخص باشد و مقدار آن اهم مهمی برای مصالحه بین پیچیدگی مدل و دقت آن است. پارامتر r_j از رابطه زیر بدست می آید :

$$r_j = \bar{r} (\bar{x}_j - \underline{x}_j) \quad (8-3)$$

که بردارهای \bar{x}_j و \underline{x}_j به ترتیب مقادیر حداکثر و حداقل مورد انتظار متغیر r_j هستند . برای \bar{r} بازه [۰/۳ ، ۰/۵] توصیه می شود.

^۱ . Generalization

بطور کلی مقادیر بزرگ r_j منجر به میانگین‌گیری^۱ و مقادیر کوچک آن به انطباق بیش از حد^۲ منتهی می‌شود یعنی مدل نهائی بجای یادگیری توزیع داده‌های ورودی-خروجی، آنها را حفظ می‌کند و قدرت تعمیم لازم را نخواهد داشت چنین مدلی در مرحله آزمایش می‌تواند خروجی متناظر با داده‌های ورودی مورد استفاده در آموزش را بدرستی تخمین بزند اما در تخمین داده‌هایی که در آموزش وجود نداشته‌اند دقت خوبی نخواهد داشت.

بعلت اینکه توزیع داده‌ها قبل از شروع آموزش مشخص نیست و مهمتر اینکه در سیستم‌های واقعی توزیع داده‌ها مرتباً در حال تغییر می‌باشد، شعاع‌های حقیقی خوشه‌ها نیز با گذشت زمان تغییر می‌کند. همچنین بعلت تفاوت ابعاد خوشه‌ها، عرضهای توابع عضویت هر متغیر خاص در قواعد مختلف با یکدیگر متفاوت است، بنابراین فرض اتخاذ شده در [۴] بهترین انتخاب ممکن نیست.

• لایه سوم (پایگاه قواعد): هر گره این لایه متناظر با قسمت مقدمه یکی از قواعد است و سطح فعالسازی^۳ یا همان نیروی آتش آن قاعده را محاسبه می‌کند بنابراین تعداد گره‌های لایه قواعد برابر با تعداد قواعد است. برای i امین قاعده خواهیم داشت:

$$\varphi_i(x) = \mu_1(x_1) * \mu_2(x_2) * \dots * \mu_n(x_n) \quad (9-3)$$

علامت (*) بیانگر t -نرم^۴ فازی است که عموماً دو عملگر ضرب و حداقل برای محاسبه آن بکار می‌روند. با استفاده از ضرب، سطح فعالسازی قاعده i ام باینصورت محاسبه خواهد شد:

$$\varphi_i(x) = \prod_{j=1}^n \mu_{ij}(x_j) = \exp\left(-\sum_{j=1}^n \frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}\right) \quad (10-3)$$

به دو علت از عملگر ضرب استفاده شده است یکی بخاطر همگرایی محاسباتی و دیگر اینکه ضرب جبری می‌تواند مجموعه‌ای از توابع عضویت گوسی یک‌بعدی را به یک تابع عضویت چندبعدی تبدیل کند.

^۱ Averaging
^۲ Over Fitting
^۳ Activation level
^۴ T-norm

- لایه چهارم (نرمالیزه کننده) : سطح فعالسازی نرمالیزه شده هر قاعده توسط یکی از گره‌های این لایه حساب می‌شود بنابراین تعداد گره‌های این لایه نیز مساوی تعداد قواعد خواهد بود . برای قاعده i ام داریم :

$$\lambda_i = \frac{\varphi_i}{\sum_{k=1}^R \varphi_k} \quad (11-3)$$

پارامتر λ_i سهم زیرسیستم i ام در خروجی نهائی مدل TS را مشخص می‌کند. از آنجا که در این لایه پارامتر قابل تنظیمی وجود ندارد غالباً لایه جداگانه‌ای برای نرمالیزه کردن در نظر نمی‌گیرند یعنی شبکه فازی- عصبی بصورت چهارلایه‌ای درمی‌آید .

- لایه پنجم (غیرفازی‌ساز یا خروجی) : هر کدام از گره‌های این لایه یکی از متغیرهای خروجی را محاسبه می‌کنند بنابراین تعداد آنها مساوی m خواهد بود . ضریب بین گره i ام لایه چهارم و گره j ام لایه پنجم y_j^i است . خروجی j ام بصورت زیر محاسبه می‌شود:

$$y_j = \sum_{i=1}^R \lambda_i y_j^i ; \quad j=1:m \quad (12-3)$$

۳-۴- شناسائی ساختار

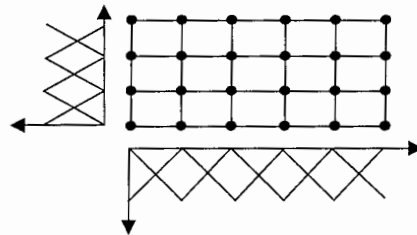
این مرحله در اکثر روشهای شناسائی فازی با یکی از دو روش بخش‌بندی شبکه‌ای^۱ و خوشه-سازی^۲ فضای ورودی-خروجی انجام می‌شود [۲۷]

• بخش‌بندی شبکه‌ای

در این روش دامنه هر کدام از متغیرهای ورودی به تعداد مشخصی بازه (مجموعه فازی) تقسیم‌بندی می‌شود. مجموعه‌های فازی هر متغیر بصورت یکسان تعریف می‌شوند و پایگاه قواعد بگونه‌ای طراحی می‌شود که بخشهای مقدمه قواعد آن شامل همه ترکیبهای ممکن از مجموعه‌های فازی ورودیها باشد تا همه فضای ورودی پوشش داده شود .

^۱ . Grid Partitioning

^۲ . Clustering



شکل ۳-۱- نمونه‌ای از بخش‌بندی شبکه‌ای فضای دوبعدی

قواعد بدست آمده با این روش متفرق و پراکنده هستند بدین معنی که به ازاء هر نمونه ورودی تنها تعداد کمی از قواعد موجود روی خروجی تاثیر می‌گذارند. مشکل دیگر روش مذکور اینست که تعداد مجموعه‌های فازی و در نتیجه تعداد قواعد تابعی نمائی از تعداد و دامنه متغیرها است که افزایش حجم پایگاه و پیچیدگی محاسباتی روش را بهمراه دارد.

• خوشه‌سازی فازی

هر ناحیه تراکم داده‌ها در فضای ورودی-خروجی، $(z = [x^T, y]^T)$ ، خوشه نامیده می‌شود. برای خوشه‌بندی فضای داده‌ها روشهای متعددی وجود دارند. روشهای خوشه‌سازی تفریقی^۱، میانگین C فازی^۲، میانگین k، نگاشتهای خودسازمانده خصوصیت^۳، روشهای یادگیری رقابتی، الگوریتم خوشه-سازی گوستافسون-کسل و روش کوه^۴ [۴] از جمله روشهایی هستند که کارآمدی آنها در آموزش پارامترهای قسمت مقدمه قواعد بصورت خارج خط (در مد یادگیری تحلیل دسته‌ای^۵) ثابت شده است. در ادامه به معرفی الگوریتم خوشه‌سازی تفریقی می‌پردازیم.

۳-۴-۱- الگوریتم خوشه‌سازی تفریقی خارج خط [۴]

در خوشه‌سازی تفریقی قابلیت هر نقطه برای تبدیل به مرکز خوشه از روی پتانسیل آن سنجیده می‌شود. همانطور که قبلاً گفته شد پتانسیل نمونه z_i ، (p_i) ، بصورت میزان فاصله فضائی بین z_i و تمام نمونه‌های دیگر تعریف می‌شود. با فرض رابطه (۳-۷) برای توابع عضویت، خواهیم داشت:

^۱ Subtractive

^۲ Fuzzy C-Means

^۳ SOM Self-Organize Feature Map

^۴ Mountain Clustering

^۵ Batch Processing

$$p_i = \frac{1}{N} \sum_{j=1}^N \exp \left(- \sum_{k=1}^{n+m} \left(\frac{(z_{ik} - z_{jk})^2}{r_k^2} \right) \right) ; i = [1, N] \quad (13-3)$$

N تعداد داده‌های آموزشی می‌باشد. از رابطه (۱۳-۳) می‌توان دید پتانسیل نقطه‌ای که توسط تعداد زیادی نقاط دیگر احاطه شده بیشتر می‌باشد و هر چه نقاط احاطه کننده نقطه مورد نظر زیادتر شوند استعداد آن برای تبدیل شدن به مرکز خوشه بیشتر می‌شود. اگر نقطه‌ای بعنوان مرکز خوشه انتخاب شود پتانسیل تمام نقاط دیگر کاهش می‌یابد که این کاهش با پتانسیل مرکز خوشه نسبت مستقیم و با فاصله از مرکز خوشه نسبت معکوس دارد. پس از این مرحله، نمونه دارای بالاترین پتانسیل بعنوان مرکز گره بعدی انتخاب می‌شود و این فرایند تا زمانی که پتانسیل کلیه نقاط از آستانه مشخصی کمتر شود ادامه می‌یابد.

روش خوشه‌سازی تفریقی شامل مراحل زیر می‌باشد:

۱- در ابتدا نقطه‌ای که بیشترین پتانسیل را دارد بعنوان مرکز اولین خوشه انتخاب می‌شود (p_1):

$$p_1^* = \max_{i=1}^N p_i \quad (14-3)$$

۲- پتانسیل تمام نقاط دیگر کاهش داده می‌شود.

$$p_i = p_i - P_1^* \exp \left(- \sum_{k=1}^{n+m} \left(\frac{(z_{ik} - z_{1k})^2}{r_{bk}^2} \right) \right) \quad i = [1, N] \quad (15-3)$$

r_{bj} ثابت مثبتی است که مقدار $1.5r$ برای آن پیشنهاد می‌شود. در حالت کلی بعد از اینکه s مرکز خوشه شناسائی شد پتانسیل نقاط دیگر بصورت زیر کاهش داده می‌شود:

$$p_i = p_i - P_s^* \exp \left(- \sum_{k=1}^{n+m} \left(\frac{(z_{ik} - z_{sk})^2}{r_{bj}^2} \right) \right) \quad (16-3)$$

۳- دو شرط مرزی آستانه پائین ($\underline{\varepsilon} p^{ref}$) و آستانه بالا ($\bar{\varepsilon} p^{ref}$) که هر دو تابعی از پتانسیل ماکزیمال (p^{ref}) هستند، انتخاب می‌شوند. اگر پتانسیل نقطه‌ای از آستانه بالا بیشتر باشد بعنوان مرکز خوشه جدید (و در نتیجه مرکز قاعده جدید) انتخاب می‌شود و اگر پتانسیل آن از آستانه پائین کمتر باشد آن نقطه نمی‌تواند مرکز خوشه باشد و الگوریتم خوشه‌سازی نیز در همین مرحله به پایان می‌رسد.

۴- اگر پتانسیل نقطه مورد نظر بین دو مقدار آستانه باشد آنگاه کمترین فاصله بین نامزد جدید مرکز خوشه و تمام مراکز خوشه‌های قبلی (δ_{min}) تعیین‌کننده خواهد بود. نامساوی زیر نوعی مصالحه بین پتانسیل نقطه مورد نظر و نزدیکی به مراکز خوشه‌های قبلی برقرار می‌کند:

$$\frac{\delta_{min}}{r} + \frac{P_k^*}{P_1^*} \geq 1 \quad (17-3)$$

در صورت برقراری نامساوی، نقطه مورد نظر بعنوان مرکز خوشه جدید انتخاب می‌شود و به مرحله ۲ برمی‌گردیم در غیر اینصورت پتانسیل آن نقطه برابر صفر قرار داده می‌شود، نقطه بعدی دارای بیشترین پتانسیل را در نظر گرفته و شرط در مورد آن بررسی می‌شود. این راهکار جهت تخمین اولیه پارامترهای مقدمه در شناسائی فازی بکار می‌رود و متکی بر این ایده است که مرکز هر خوشه، بیانگر مشخصه‌ای از رفتار سیستم می‌باشد. مراکز خوشه‌های بدست آمده بعنوان پارامترهای قسمتهای مقدمه مورد استفاده قرار می‌گیرند که بیانگر نقاط کانونی قواعد مدل نیز هستند.

۲-۴-۳- خوشه‌سازی تفریقی برخط

در روشهای شناسائی سیستم برخط، داده‌های آموزشی بجای اینکه یک مجموعه ثابت باشند بطور پیوسته تغییر می‌کنند. بعضی از داده‌های جدید اطلاعات موجود در داده‌های قبلی را تقویت و تأیید می‌کنند و داده‌های دیگر حاوی اطلاعات جدیدی مانند تغییر در شرایط کاری، ایجاد خطا یا تغییر دینامیک سیستم هستند که در صورت کافی بودن این اطلاعات جدید، قاعده تازه‌ای به پایگاه قواعد اضافه می‌شود یا اینکه یکی از قواعد موجود بهبود می‌یابد.

یادگیری برخط مدل‌های eTS^۱ شامل خوشه‌سازی برخط با فرض تغییرات آهسته پایگاه قواعد و روش کمترین مربعات بازگشتی (وزن‌دار) بهبودیافته انجام می‌شود. بعلت تکامل ساختار مدل انتظار می‌رود که تعداد قواعد فازی افزایش یابد اما این امر نگران‌کننده نیست چون پتانسیل رابطه معکوسی با تعداد نقاط دارد.

روش خوشه‌سازی برخط با انتخاب اولین داده بعنوان نقطه کانونی خوشه اول آغاز می‌شود. با استفاده از یکی از انواع توابع عضویت (مثلاً توابع گوسی) و مختصات این نقطه، قسمت مقدمه اولین قاعده فازی را تشکیل می‌دهیم. بجای توابع گوسی می‌توانیم هر کدام از انواع مختلف توابع عضویت را بکار برده و پتانسیل آنرا برابر یک قرار دهیم. با شروع از داده بعدی، پتانسیل داده‌های جدید بطور بازگشتی محاسبه می‌شود که بدین منظور از تابع نوع کوشی مرتبه یک استفاده می‌شود:

$$p_k(z_k) = \frac{1}{1 + \frac{1}{(k-1)} \sum_{l=1}^{k-1} \sum_{j=1}^{n+m} (d_{lk}^j)^2} ; k = 2, 3, \dots \quad (18-3)$$

$p_k(z_k)$ پتانسیل داده z_k محاسبه شده در زمان k است و $d_{lk}^j = z_l^j - z_k^j$ تصویر فاصله بین دو داده $(z_l^j$ و $z_k^j)$ روی محور z^j را نشان می‌دهد. (برای $j = 1, \dots, n$ روی محور x^j و برای $j = n, \dots, n+m$ روی محور y).

این تابع یکنواخت دارای تناسب معکوس با فاصله بوده و امکان محاسبه آن بصورت بازگشتی وجود دارد که این خاصیت جهت پیاده‌سازی الگوریتم یادگیری بصورت برخط اهمیت زیادی دارد. بعد از حاضر شدن هر داده جدید بصورت برخط، تمام پتانسیل‌های موجود را بروز می‌کنیم. پتانسیل داده جدید مطابق رابطه زیر بصورت بازگشتی محاسبه می‌شود:

$$p_k(z_k) = \frac{(k-1)}{(k-1)(\nu_k + 1) + \sigma_k - 2\nu_k} \quad (19-3)$$

در این رابطه داریم:

^۱ Evolving Takagi-Sugeno

$$\omega_k = \sum_{j=1}^{n+m} (z_k^j)^2; \quad \tau_k = \sum_{l=1}^{k-1} \sum_{j=1}^{n+m} (z_l^j)^2$$

$$v_k = \sum_{j=1}^{n+m} z_k^j \beta_k^j; \quad \beta_k^j = \sum_{l=1}^{k-1} z_l^j$$

پارامترهای v_k و u_k از روی داده جاری (z_k) محاسبه می‌شوند در حالیکه β_k^j و σ_k بصورت بازگشتی مطابق روابط زیر بدست می‌آیند :

$$\sigma_k = \sigma_{k-1} + \sum_{j=1}^{n+1} (Z_{k-1}^j)^2 \quad (20-3)$$

$$\beta_k^j = \beta_{k-1}^j + Z_{k-1}^j \quad (21-3)$$

بعد از آماده شدن هر داده جدید در مد برخط ، پتانسیل‌های مراکز خوشه‌ها $(c_l; l=[1, R])$ که متناظر با نقاط کانونی قواعد موجود $(x_l^*, l=[1, R])$ هستند تحت تاثیر قرار می‌گیرند. علت این تاثیر ، وابستگی تعریف ارائه شده برای پتانسیل هر نقطه به فاصله آن نقطه از تمام نقاط دیگر (شامل نقطه جدید) می‌باشد. رابطه بازگشتی برای بروز کردن پتانسیلهای نقاط کانونی خوشه‌های موجود را می‌توان به آسانی از رابطه (۱۹-۳) بدست آورد :

$$p_k(c_l) = \frac{(k-1)p_{k-1}(c_l)}{[k-2 + p_{k-1}(c_l) + p_{k-1}(c_l) \sum_{j=1}^{n+m} (d_{k(k-1)}^j)^2]} \quad (22-3)$$

که $p_k(c_l)$ پتانسیل مرکز خوشه (نمونه اصلی قاعده am) در زمان k می‌باشد. پتانسیلهای نقاط جدید با پتانسیل بروز شده مراکز خوشه‌های موجود مقایسه می‌شود.

اگر پتانسیل داده جدید باندازه کافی بالا باشد (معیار تشخیص بالا بودن پتانسیل در ادامه معرفی خواهد شد) این داده بعنوان یک مرکز جدید پذیرفته شده و قاعده جدیدی ایجاد می‌شود که نقطه کانونی بر اساس تصویر^۱ این مرکز جدید روی محور x $(x_R^* = x_k; R = R+1)$ بدست می‌آید. در این حالت نقطه جدید توصیف بهتری از سیستم داشته و قدرت خلاصه‌کنندگی آن از تمام داده-های قبلی بیشتر است. تذکر این نکته لازمست که تحقق شرط بالا بودن پتانسیل داده جدید مشکل می‌باشد چون با افزایش تعداد داده‌ها غالباً تمرکز آنها کمتر می‌شود (به استثناء برخی نواحی مهم

^۱.Projection

جدید که در فضای کاری ظاهر شده و بوجود آمدن نواحی کاری یا شرایط جدیدی را نشان می‌دهند (در این وضعیتها قاعده جدیدی ساخته می‌شود و در عین حال داده‌های پرت بطور خودکار پذیرفته نمی‌شوند چون فاصله‌شان از بقیه داده‌ها زیاد بوده و در نتیجه پتانسیل آنها بمیزان قابل توجهی پائین می‌باشد. این خاصیت روش ارائه شده جهت کاربرد در مسائل تشخیص خطا بسیار امیدوارکننده می‌باشد.

اگر پتانسیل داده جدید از پتانسیل تمام مراکز قبلی بیشتر باشد و اضافه بر آن ، داده جدید به یکی از مراکز قبلی خیلی نزدیک باشد (معیار تشخیص کوچک بودن فاصله مذکور نیز در ادامه معرفی خواهد شد) آنگاه داده جدید (z_k) جایگزین این مرکز می‌شود ($z_c = z_k$). همچنین در این پایان-نامه معیاری برای حذف قواعد اضافی معرفی شده است. مکانیزم اضافه کردن قاعده جدید ، جایگزینی قواعد موجود و حذف قواعد اضافی اصلاح^۱ پایگاه نامیده می‌شود.

۳-۵- آموزش پارامترهای زیرمدلهای خطی (شناسائی پارامتر)

با ثابت ماندن پارامترهای قسمت مقدمه می‌توان تخمین پارامترهای مدل‌های خطی قسمت نتیجه را به مساله کمترین مربعات تبدیل کرد. اینکار با صرفنظر از عملگر مجموع در رابطه (۳-۱۲) و جایگزینی آن با عبارت برداری معادل انجام می‌شود :

$$\hat{y}_{k+1} = \psi_k^T \theta_k \quad (۳-۲۳)$$

θ برداری مرکب از پارامترهای مدل‌های خطی و ψ بردار ورودیهاست که بوسیله سطوح فعالسازی نرمالیزه شده قواعد وزن‌دهی شده است :

$$\theta = [\pi_1^T, \pi_2^T, \dots, \pi_R^T]^T \quad (۳-۲۴)$$

$$\psi = [\lambda_1 x_e^T, \lambda_2 x_e^T, \dots, \lambda_R x_e^T]^T \quad (۳-۲۵)$$

^۱ Modification

برای یک مجموعه داده ورودی-خروجی $\{(x_k^T, y_k), k = [1, N]\}$ بردار پارامترهای خطی θ باید تابع هدف کلی^۱ زیر را حداقل کند:

$$J_G = \sum_{k=1}^N (y_k - \psi_k^T \theta)^T \quad (26-3)$$

تابع هدف (۲۶-۳) را می‌توان بفرم برداری زیر نوشت :

$$J_G = (Y - \Psi^T \theta)^T (Y - \Psi^T \theta) \quad (27-3)$$

حداقل شدن تابع هدف (۲۷-۳) تضمین نمی‌کند که رفتار محلی زیرمدلهای مدل TS مناسب باشد. زیرمدلهای بهینه محلی از روی تابع هدف وزن داده شده بصورت محلی بدست می‌آیند :

$$J_L = \sum_{i=1}^R (Y - X^T \pi_i)^T \Lambda_i (Y - X^T \pi_i) \quad (28-3)$$

ماتریس $X \in R^{N \times (n+1)}$ از روی x_{ek}^T ساخته می‌شود و Λ_i نیز ماتریسی قطری است که عناصر روی قطر اصلی آن $\lambda_i(x_k)$ می‌باشند. مجموعه جوابهای متعلق به توابع هزینه جداگانه (بردارهای π_i) را می‌توان به روش کمترین مربعات بازگشتی وزن دار^۲ (WRLS) محاسبه کرد.

در برخورد با مساله افزایش حجم داده‌های آموزشی برای حالات بهینه کلی و بهینه محلی بترتیب از الگوریتمهای RLS و WRLS استفاده می‌شود. در این روشها فرض بر اینست که پایگاه قواعد ثابت است (بعبارت دیگر پارامترهای قسمت مقدمه تغییر نمی‌کنند). با پذیرش این فرض ، مسائل بهینه-سازی ۳-۴۷ ، ۳-۴۸ و ۳-۵۱ از نظر پارامتر خطی می‌باشند . همچنین در روش eTS فرض می‌شود تکامل پایگاه قواعد بصورت تدریجی انجام می‌شود بنابراین تعداد قواعد و پارامترهای قسمت مقدمه تغییر خواهند کرد اگر چه این تغییرات غالباً بمیزان قابل توجهی نسبت به پله زمانی (تغییرات در مجموعه داده‌ها) کمتر اتفاق می‌افتند .

بعلت انجام تکامل ذکر شده ، سطح فعالسازی نرمالیزه شده قواعد (λ_i) تغییر خواهد کرد و از آنجا که این تغییرات روی همه داده‌ها (شامل داده‌های جمع‌آوری شده قبل از زمان هر تغییر) تاثیر

^۱ Global

^۲ Weighted Recursive Least Square

می‌گذارد کاربرد مستقیم RLS یا WRLS صحیح نیست. با اضافه شدن هر قاعده به پایگاه قواعد یا حذف آن، ماتریسهای کواریانس فیلتر کالمن (RLS) باید بازنشانی شوند و پارامترهای آن نیز مقاداردهی اولیه شوند. تخمین پارامترها و ماتریسهای کواریانس قاعده جدید (R+1) ام از روی میانگین وزن دار پارامترها و ماتریسهای کواریانس متناظر (R) قاعده قبلی بدست می‌آید.

۳-۵-۱- تخمین کلی پارامترها

مدل eTS جهت پیش‌بینی برخط خروجی بر اساس ورودیهای قبلی مورد استفاده قرار می‌گیرد:

$$\hat{y}_{k+1} = \psi_k^T \hat{\theta}_k \quad ; \quad k=2,3,\dots \quad (29-3)$$

روش فیلتر کالمن بصورت زیر اعمال می‌شود:

$$\hat{\theta}_k = \hat{\theta}_{k-1} + C_k \psi_{k-1} (y_k - \psi_{k-1}^T \hat{\theta}_{k-1}) \quad (30-3)$$

$$C_k = C_{k-1} - \frac{C_{k-1} \psi_{k-1} \psi_{k-1}^T C_{k-1}}{1 + \psi_{k-1}^T C_{k-1} \psi_{k-1}} \quad (31-3)$$

در اینجا شرایط اولیه بصورت $\hat{\theta}_0 = 0$ و $C_0 = \Omega I$ هستند که Ω یک مقدار مثبت بزرگ و C ماتریس کوواریانس با ابعاد $R(n+1) * R(n+1)$ می‌باشند. $\hat{\theta}_k$ تخمینی از پارامترها بر اساس k نمونه داده می‌باشد.

وقتی قاعده جدیدی به پایگاه قواعد اضافه می‌شود فیلتر کالمن بصورت زیر بازنشانی می‌شود:

۱) پارامترهای قاعده جدید بصورت میانگین وزن دار پارامترهای دیگر قواعد محاسبه می‌شوند و از سطوح فعالسازی نرمالیزه شده قواعد موجود (λ_i) بعنوان ضرایب وزنی استفاده می‌شود. پارامترهای قواعد دیگر از مرحله قبلی تکرار گرفته می‌شوند:

$$\hat{\theta}_k = [\hat{\pi}_{1(k-1)}^T, \hat{\pi}_{2(k-1)}^T, \dots, \hat{\pi}_{R(k-1)}^T, \hat{\pi}_{(R+1)k}^T]^T \quad (32-3)$$

$$\hat{\pi}_{(R+1)k} = \sum_i^R \lambda_i \hat{\pi}_{i(k-1)} \quad (33-3)$$

۲) ماتریسهای کوواریانس بصورت زیر بازنشانی می‌شوند:

$$C_k = \begin{bmatrix} \rho \zeta_{11} & \dots & \rho \zeta_{1R(n+m)} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho \zeta_{R(n+m)1} & \dots & \rho \zeta_{R(n+m)R(n+m)} & 0 & \dots & 0 \\ 0 & 0 & 0 & \Omega & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & \Omega \end{bmatrix} \quad (34-3)$$

که ζ_{ij} عضوی از ماتریس کواریانس است ($i = [1, R(n+m)]; j = [1, R(n+m)]$) و $\rho = ((R^2 + 1)/R^2)$ یک ثابت می‌باشد. وقتی یک قاعده توسط قاعده دیگر جایگزین می‌شود و پارامترهای مقدمه هر دو قاعده نزدیک به یکدیگر باشند آنگاه پارامترها و ماتریسهای کواریانس از تکرار قبلی گرفته می‌شوند.

۳-۵-۲- تخمین محلی پارامترها

این تخمین بر روش WRLS مبتنی می‌باشد:

$$\hat{\pi}_{ik} = \hat{\pi}_{i(k-1)} + C_{ik} x_{ek-1} \lambda_i(x_{k-1}) (y_k - x_{ek-1}^T \hat{\pi}_{ik-1}) \quad ; \quad k = 2, 3, \dots \quad (35-3)$$

$$C_{ik} = C_{i(k-1)} - \frac{\lambda_i(x_{k-1}) C_{i(k-1)} x_{ek-1} x_{ek-1}^T C_{i(k-1)}}{1 + \lambda_i(x_{k-1}) x_{ek-1}^T C_{i(k-1)} x_{ek-1}} \quad (36-3)$$

شرایط اولیه نیز بصورت زیر هستند:

$$\hat{\pi}_1 = 0 \quad ; \quad C_{i1} = \Omega I \quad (37-3)$$

در این حالت ماتریسهای کواریانس برای هر قاعده مجزا بوده و ابعاد کوچکتری دارند. بطوریکه داریم: $(C_{ik} \in R^{(n+1) \times (n+1)}; i = [1, R])$. پارامترهای قاعده جدید بصورت میانگین وزن دار پارامترهای قواعد قبلی و از رابطه (۳۳-۳) محاسبه می‌شوند و پارامترهای R قاعده قبلی هم بدون تغییر از تکرار قبلی گرفته می‌شوند:

$$\pi_{ik} = \pi_{i(k-1)} \quad ; \quad i = [1, R] \quad (38-3)$$

وقتی قاعده جدیدی جایگزین قاعده دیگر می‌شود و پارامترهای مقدمه (مراکز) دو قاعده به یکدیگر نزدیکی زیادی داشته باشند آنگاه پارامترهای تمام قواعد بدون تغییر باقی می‌مانند:

$$\pi_{ik} = \pi_{i(k-1)} \quad ; \quad i = [1, R] \quad (39-3)$$

پارامترهای R قاعده قبلی بدون تغییر از تکرار قبلی گرفته شده و مقداردهی اولیه ماتریس کواریانس قاعده جدید بصورت زیر انجام می‌شود:

$$C_{(R+1)k} = \Omega I \quad ; \quad C_{ik} = C_{i(k-1)} \quad ; \quad i = [1, R] \quad (40-3)$$

۳-۶- اصلاح شعاعهای تاثیر قواعد

در اکثریت قریب باتفاق مراجع معدود مربوط به شناسائی سیستم برخط فازی، شعاعهای تاثیر قاعده جدید بدلتخواه انتخاب می‌شود که با توجه به کمبود اطلاعات در مورد شعاع واقعی خوشه جدید انجام این انتخاب با دقت مناسب مشکل است. در اینجا پیشنهاد می‌شود با ایجاد هر قاعده جدید شعاعهای تاثیر آن بصورت زیر تعیین شوند:

$$\sigma_{(R+1)j} = \frac{\max\{|z_j - c_{aj}|, |z_j - c_{bj}|\}}{D} * n \quad (41-3)$$

در صورتیکه عرض توابع عضویت موجود در هر قاعده برای متغیرهای مختلف یکسان در نظر گرفته شود (استفاده از σ_i بجای σ_{ij})، آنگاه داریم:

$$\sigma_{R+1} = \frac{\max(\|c_{R+1} - c_{n1}\|, \|c_{R+1} - c_{n2}\|)}{D} \quad (42-3)$$

c_{n1} و c_{n2} مراکز هستند که از بقیه مراکز موجود به مرکز c_i نزدیکترند و D نیز یک عدد ثابت می‌باشد که معمولاً در بازه $[1/5 - 2/5]$ در نظر گرفته می‌شود. از آنجا که شعاع تاثیر خوشه‌های موجود در هر ناحیه به تعداد آنها نیز بستگی دارد بنابراین در تعیین D تعداد قواعد را نیز دخالت می‌دهیم. رابطه پیشنهادی بصورت زیر می‌باشد:

$$D = \sqrt{\ln\left(\frac{1}{F_{gen}}\right)} \quad (43-3)$$

$$F_{gen} = \min\left[F_{min} * \rho^{-R}, F_{max}\right]$$

در رابطه بالا ρ در محدوده $(0,1)$ بوده و ثابت تنزل^۱ نامیده می‌شود. F_{\max} و F_{\min} نیز دو ثابت می‌باشند که همانند ρ باید در ابتدای الگوریتم انتخاب شوند.

بعد از تخصیص مراکز خوشه‌ها و شعاعهای تاثیر آنها، شبکه فازی عصبی بصورت یک شبکه پیش‌خور^۲ خطی درمی‌آید بنابراین برای اصلاح پارامترهای قسمت نتیجه احتیاجی به الگوریتمهای مبتنی بر پس‌انتشار نمی‌باشد. در مرحله اصلاح پارامترهای قسمت نتیجه، پارامترهای قسمت مقدمه (مراکز خوشه‌ها و شعاع تاثیر آنها) ثابت فرض می‌شوند.

یکی از الگوریتمهای پرکاربرد در مسائل بهینه‌سازی الگوریتم کاهش گرادیان است که بیشترین کاربرد آن در شناسائی پارامتر مربوط به روشهای شناسائی خارج‌خط می‌باشد و کاربرد آن در روشهای شناسائی برخط بسیار محدود بوده و مختص به ساختارهای خاصی از مدلها بوده است (مدل FIR در مرجع [۶]). در اینجا الگوریتم مذکور را برای اصلاح پارامترهای شبکه فازی-عصبی عمومی معرفی شده در بخش (۳-۳-۲) گسترش داده‌ایم. ابتدا خطا در لحظه k را بصورت زیر معرفی می‌کنیم:

$$E = \frac{1}{2} \sum_{r=1}^m e_r^2(k) = \frac{1}{2} \sum_{r=1}^m [y_r(k) - \hat{y}_r(k)]^2 \quad (۴۴-۳)$$

همچنین با توجه به ساختار شبکه فازی-عصبی، تعاریف زیر را ارائه می‌دهیم:

$$\delta_r^{(4)} = \frac{\partial E}{\partial \hat{y}_r} ; \quad \delta_i^{(3)} = \frac{\partial E}{\partial \varphi_i} ; \quad \delta_{ij}^{(2)} = \frac{\partial E}{\partial \mu_{ij}} \quad (۴۵-۳)$$

برای محاسبه جملات پس‌انتشار خطا خواهیم داشت:

$$\delta_r^{(4)} = \frac{\partial E}{\partial \hat{y}_r} = -(y_r(k) - \hat{y}_r(k)) \quad (۴۶-۳)$$

$$\delta_i^{(3)} = \frac{\partial E}{\partial \varphi_i} = \sum_{r=1}^m \frac{\partial E}{\partial \hat{y}_r} \frac{\partial \hat{y}_r}{\partial \varphi_i} = \sum_{r=1}^m \delta_r^{(4)} \frac{\partial \hat{y}_r}{\partial \varphi_i} \quad (۴۷-۳)$$

که داریم

^۱. Decay Constant

^۲. Feed Forward

$$\frac{\partial \hat{y}_r}{\partial \varphi_i} = \frac{\partial}{\partial \varphi_i} \left[\frac{\sum_{h=1}^R \varphi_h \hat{y}_r^h}{\sum_{h=1}^R \varphi_h} \right] = \frac{\sum_{h=1}^R \varphi_h (\hat{y}_r^i - \hat{y}_r^h)}{\left(\sum_{h=1}^R \varphi_h \right)^2} = \frac{\hat{y}_r^i - \hat{y}_r}{\sum_{h=1}^R \varphi_h} \quad (48-3)$$

در نتیجه جمله خطا برای لایه سوم باین شکل خلاصه خواهد شد :

$$\delta_i^{(3)} = \sum_{r=1}^m \delta_r^{(4)} \frac{\hat{y}_r^i - \hat{y}_r}{\sum_{h=1}^R \varphi_h} \quad (49-3)$$

برای لایه دوم خواهیم داشت :

$$\delta_{ij}^{(2)} = \frac{\partial E}{\partial \mu_{ij}} = \sum_{h=1}^R \frac{\partial E}{\partial \varphi_h} \frac{\partial \varphi_h}{\partial \mu_{ij}} = \frac{\partial E}{\partial \varphi_i} \frac{\partial \varphi_i}{\partial \mu_{ij}} = \delta_i^{(3)} \frac{\partial \varphi_i}{\partial \mu_{ij}} \quad (50-3)$$

$$\varphi_i = \prod_{j=1}^n \mu_{ij}(x_j) \Rightarrow \frac{\partial \varphi_i}{\partial \mu_{ij}} = \prod_{l=1, l \neq j}^n \mu_{il}(x_l) \quad (51-3)$$

و در نهایت داریم :

$$\delta_{ij}^{(2)} = \delta_i^{(3)} \prod_{l=1, l \neq j}^n \mu_{il}(x_l) \quad (52-3)$$

پارامترهای قابل تنظیم توابع عضویت شامل c_{ij} و σ_{ij} (یا σ_i) می‌باشند فرمولهای اصلاح این

پارامترها بصورت زیر بدست می‌آیند :

$$\Delta c_{ij} = -\eta \frac{\partial E}{\partial c_{ij}} = -lr_c \left(\frac{\partial E}{\partial \mu_{ij}} \right) \frac{\partial \mu_{ij}}{\partial c_{ij}} = -lr_c \delta_{ij}^{(2)} \frac{\partial \mu_{ij}}{\partial c_{ij}} = 2lr_c \delta_{ij}^{(2)} \mu_{ij}(x_j) \left(\frac{x_i - c_{ij}}{\sigma_{ij}^2} \right) \quad (53-3)$$

با توجه به اینکه

$$\delta_{ij}^{(2)} \mu_{ij}(x_j) = \delta_i^{(3)} \varphi_i \Rightarrow \Delta c_{ij} = 2lr_c \delta_i^{(3)} \varphi_i \left(\frac{x_i - c_{ij}}{\sigma_{ij}^2} \right)$$

$$\begin{aligned} \Delta c_{ij} &= 2lr_c \left(\sum_{r=1}^m \delta_r^{(4)} \frac{\hat{y}_r^i - \hat{y}_r}{\sum_{h=1}^R \lambda_h} \right) \varphi_i \left(\frac{x_i - c_{ij}}{\sigma_{ij}^2} \right) \\ &= 2lr_c \left(\frac{\varphi_i}{\sum_{h=1}^R \varphi_h} \right) \left[\sum_{r=1}^m (y_r - \hat{y}_r) (\hat{y}_r^i - \hat{y}_r) \right] \left(\frac{x_i - c_{ij}}{\sigma_{ij}^2} \right) \end{aligned} \quad (54-3)$$

در نتیجه برای اصلاح c_{ij} خواهیم داشت :

$$c_{ij} = c_{ij} + \Delta c_{ij} \quad (55-3)$$

$$\Delta c_{ij} = 2lr_c \lambda_i \left(\sum_{r=1}^m (y_r - \hat{y}_r) (\hat{y}_r^i - \hat{y}_r) \right) \left(\frac{x_i - c_{ij}}{\sigma_{ij}^2} \right) \quad (56-3)$$

و با انجام عملیات مشابه، روابط زیر برای اصلاح σ_{ij} بدست می‌آیند :

$$\sigma_{ij} = \sigma_{ij} + \Delta \sigma_{ij} \quad (57-3)$$

$$\Delta \sigma_{ij} = 2lr_\sigma \lambda_i \left(\sum_{r=1}^m (y_r - \hat{y}_r) (\hat{y}_r^i - \hat{y}_r) \right) \left(\frac{(x_j - c_{ij})^2}{\sigma_{ij}^3} \right) \quad (58-3)$$

که پارامترهای lr_c و lr_σ بترتیب سرعت‌های یادگیری مربوط به پارامترهای c_{ij} و σ_{ij} می‌باشند. در صورتیکه عرض توابع عضویت موجود در هر قاعده برای متغیرهای مختلف یکسان در نظر گرفته شود (استفاده از σ_i بجای σ_{ij})، آنگاه داریم :

$$\Delta c_{ij} = -\frac{\partial E}{\partial c_i} = 2lr_c \lambda_i \left(\sum_{r=1}^m (y_r - \hat{y}_r) (\hat{y}_r^i - \hat{y}_r) \right) \left(\frac{x_i - c_{ij}}{\sigma_i^2} \right) \quad (59-3)$$

اصلاح شعاع تاثیر قواعد نیز باینصورت انجام می‌شود :

$$\sigma_i = \sigma_i + \Delta \sigma_i \quad (60-3)$$

$$\Delta \sigma_i = -\frac{\partial E}{\partial \sigma_i} = 2lr_\sigma \lambda_i \left(\sum_{r=1}^m (y_r - \hat{y}_r) (\hat{y}_r^i - \hat{y}_r) \right) \frac{\|x - c_i\|^2}{\sigma_i^3} \quad (61-3)$$

۳-۷- الگوریتم تکامل مدل‌های eTS

روش بازگشتی آموزش برخط مدل‌های eTS شامل مراحل زیر می‌شود:

۱. مرحله اول: تعیین ساختار اولیه پایگاه قواعد

در این مرحله ساختار اولیه پایگاه (قسمت مقدمه قواعد) مشخص می‌شود. ممکن است پایگاه

قواعد در ابتدا تنها شامل یک قاعده باشد که می‌توان آنرا از اولین داده بدست آورد:

$$k=1; R=1; x_1^* = x_k; p_1(c_1) = 1 \quad (۳-۶۲)$$

$$\theta_1 = \pi_1 = 0; C_1 = \Omega I$$

که C_1 مرکز اولین خوشه و x_1^* نقطه کانونی اولین قاعده است که در حقیقت تصویر C_1 روی محور X می‌باشد. در حالت کلی ساختار اولیه پایگاه قواعد را می‌توان با توجه به دانش اولیه موجود در مورد سیستم تعیین کرد که بدین منظور از روشهای شناسائی خارج خط استفاده می‌شود. در این حالت داریم:

$$R=R^{ini}; p_1(c_i) = 1; i = [1, R^{ini}] \quad (۳-۶۳)$$

R^{ini} تعداد قواعد اولیه تعیین شده بروشهای خارج خط می‌باشد.

۲. مرحله دوم: گرفتن داده بعدی و اصلاح مراکز و شعاعهای توابع عضویت (شروع حلقه)

در پله زمانی بعدی ($k = k + 1$)، داده جدید (z_k) گرفته می‌شود. $y(k)$ (خروجی واقعی در لحظه k) در این مرحله معلوم است و $\hat{y}(k)$ (خروجی تخمینی متناظر با لحظه k) نیز از مرحله قبل در دسترس است بنابراین می‌توانیم بوسیله روابط (۴-۵۵) تا (۴-۵۸) یا روابط (۴-۵۹) الی (۴-۶۱) مراکز و عرضهای توابع عضویت موجود را بروزرسانی کنیم.

۳. مرحله سوم: محاسبه پتانسیل داده جدید

پتانسیل داده جدید بصورت بازگشتی از رابطه (۳-۱۹) محاسبه می‌شود. با توجه به اینکه تعداد داده‌ها در روش برخط بطور پیوسته افزایش یافته و رابطه (۳-۱۹) بطور معمول از روی ماتریسهای بزرگی محاسبه می‌شود، محاسبه پارامترهای β_k^j و σ_k بصورت بازگشتی باعث صرفه‌جویی قابل توجهی در زمان و حجم محاسبات می‌شود.

۴. مرحله چهارم: بروزرسانی پتانسیل مراکز خوشه‌های موجود

در این مرحله پتانسیل‌های نقاط کانونی (مراکز) قبلی را با در نظر گرفتن تاثیر داده‌های جدید بصورت بازگشتی محاسبه می‌شوند (رابطه ۳-۲۲)

۵. مرحله پنجم: تصمیم‌گیری در مورد اضافه کردن قاعده جدید یا اصلاح قواعد

پتانسیل داده جدید با پتانسیل‌های بروز شده مراکز قبلی مقایسه می‌شود و تصمیم‌گیری برای بهبود یا افزایش درجه^۱ پایگاه قواعد این مقایسه و بصورت زیر اتخاذ می‌گردد:

الف) اگر فقط شرط بروزرسانی پایگاه قواعد برآورده شود (یعنی یکی از نامساویهای ستون اول جدول برقرار باشد) آنگاه یک قاعده جدید به پایگاه قواعد اضافه می‌شود که داده جدید نمونه اصلی نقطه کانونی آن خواهد بود (مرکز قاعده)

$$R = R + 1 \quad ; \quad x_R^* = x_k \quad ; \quad c_R = z_k \quad ; \quad p_k(c_R) = P_k(z_k) \quad (۳-۶۴)$$

پارامترهای قسمت نتیجه و ماتریسهای کواریانس برای حالات بهینه کلی و بهینه محلی بترتیب با روابط (۳-۳۳ تا ۳-۳۴) و (۳-۳۸ تا ۳-۴۱) بازنشانی می‌شوند.

ب) اگر شرط بروزرسانی پایگاه قواعد برآورده شود (یکی از نامساویهای ستون اول جدول ۳-۱ برقرار باشد) همچنین داده جدید باندازه کافی به مرکز شماره j نزدیک باشد (یعنی نامساوی موجود در ستون دوم جدول ۳-۱ که با نامساوی برقرار شده از ستون اول متناظر است نیز برقرار شود) آنگاه داده جدید (z_k) جایگزین این مرکز می‌شود. در این حالت خواهیم داشت:

$$c_j = \arg \min_{i=1}^R (\|z_i - c_i\|) : \quad x_j^* = x_k ; \quad p_k(c_j) = P_k(z_k) \quad (۳-۶۵)$$

پارامترهای مقدمه و ماتریسهای کواریانس قاعده جایگزین همان پارامترها و ماتریسهای قاعده خارج شده از پایگاه می‌باشند:

$$\hat{\pi}_k = \pi_k^* \quad ; \quad C_k = C_k^* \quad (۳-۶۶)$$

^۱. Upgrade

با جایگزین شدن یک قاعده توسط قاعده دیگر، از آنجا که مرکز جدید به مرکز جایگزین شده نزدیک است می‌توانیم از اغتشاش وارده به فیلتر کالمن در اثر این تغییر صرفنظر کنیم چون فیلتر کالمن می‌تواند با استفاده از تخمینهای جاری پارامترها و ماتریسهای کواریانس بر اغتشاش مذکور غلبه کند.

۶. مرحله ششم: حذف قواعد اضافی

از آنجا که الگوریتم مدل‌سازی معرفی شده بیشتر بمنظور کاربرد در سیستمهای دارای دینامیک متغیر با زمان معرفی شده است و در اینگونه سیستمها توزیع داده‌ها مرتباً در حال تغییر می‌باشد بنابراین ممکن است بعضی از قواعد فعال در لحظه حاضر بعد از گذشت مدت زمانی تبدیل به قواعد غیرفعالی شوند که اثر ناچیزی در تعیین خروجی دارند. وجود اینگونه قواعد تنها باعث افزایش حجم محاسبات می‌شود بنابراین باید قواعد اضافی تشخیص داده شده و حذف گردند. در این بخش به معرفی معیاری برای تشخیص اینگونه قواعد می‌پردازیم. اگر چه اعمال این معیار به خودی خود باعث افزایش حجم محاسبات می‌شود اما در بخش شبیه‌سازی خواهیم دید که با توجه به کاهش قابل توجه تعداد قواعد و راهکارهای معرفی شده برای کاهش محاسبات اعمال معیار حذف، استفاده از این معیار کاملاً مقرون به صرفه می‌باشد.

فرض کنیم که در لحظه k پایگاه قواعد دارای R' قاعده باشد. پارامترهای d_{\min} و p_{\max} را بصورت زیر معرفی می‌کنیم:

$$d_{\min} = \min \left\{ \exp \left(- \frac{2 \|c_i - c_l\|}{\|\sigma_i\| + \|\sigma_l\|} \right) \right\} ; \quad i=1:R'-1 \quad ; \quad l=2:R' \quad (۶۷-۳)$$

$$p_{\max} = \max_{i=1}^R p_k(c_i) \quad (۶۸-۳)$$

همچنین فرض می‌کنیم c_a و c_b مراکز دو خوشه باشند که نزدیکترین فاصله را با یکدیگر دارند و داشته باشیم $p_k(c_a) < p_k(c_b)$ ، اگر رابطه زیر برقرار باشد:

$$\frac{d_{\min}}{\|\sigma_a\|} + \frac{p_k(c_a)}{p_{\max}} < 1 \quad (۶۹-۳)$$

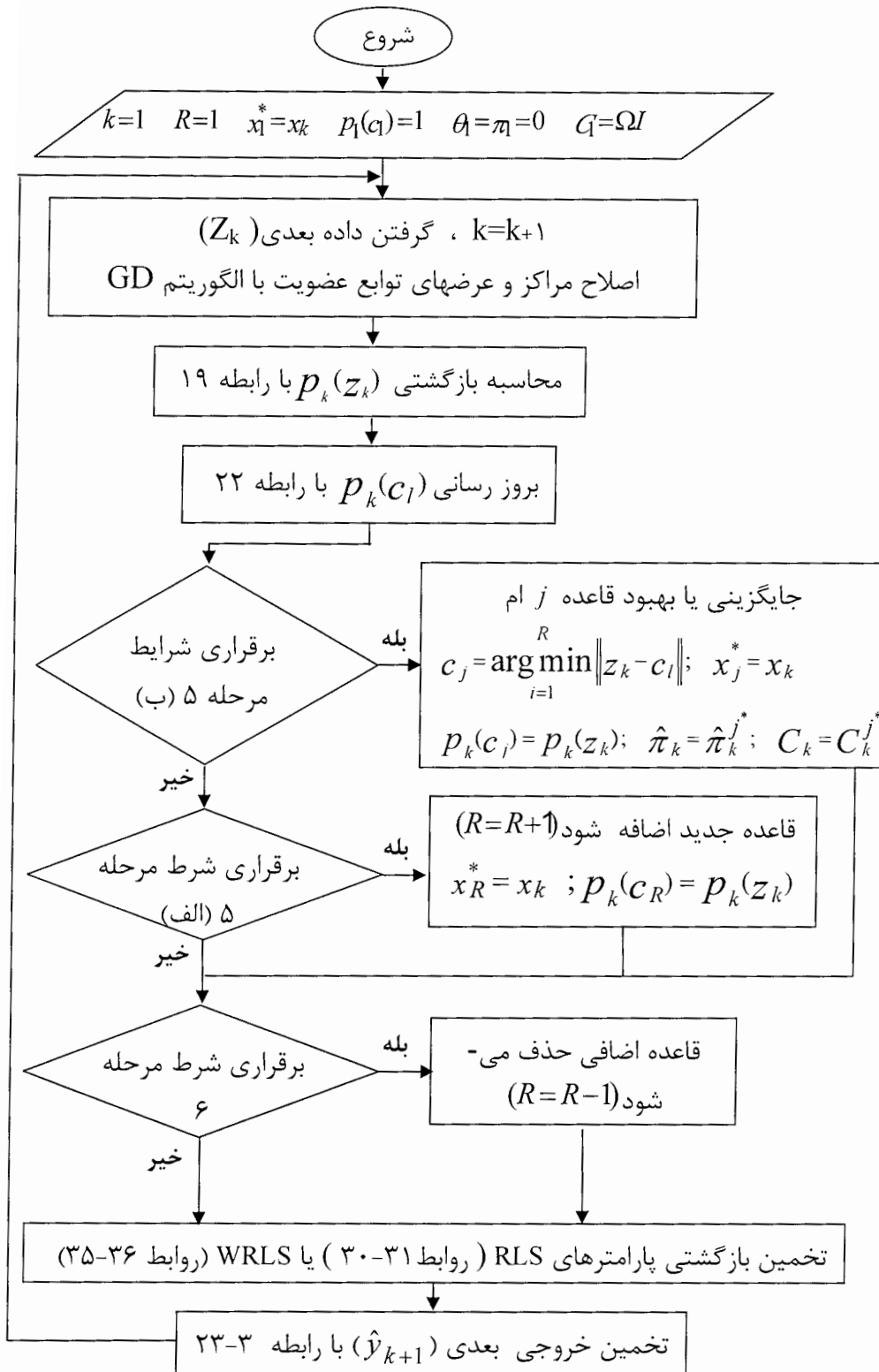
آنگاه قاعده a از پایگاه قواعد حذف خواهد شد.

۷. مرحله هفتم: بروزرسانی پارامترهای قسمت نتیجه قواعد

در مرحله ۷ پارامترهای قسمت نتیجه را بصورت بازگشتی بروز می‌کنیم. اینکار برای پارامترهای بهینه کلی با روش RLS (روابط ۳-۳۰ و ۳-۳۱) انجام می‌شود و پارامترهای بهینه محلی با استفاده از روش WRLS (روابط ۳-۳۵ و ۳-۳۶) بروز می‌شوند.

۷. مرحله هشتم: تخمین خروجی مرحله بعد و تکرار حلقه شناسایی

در این مرحله خروجی سیستم در پله زمانی بعدی ($k+1$) با استفاده از مدل eTS بدست آمده (رابطه ۳-۲۳) محاسبه می‌شود. الگوریتم شناسایی با برگشت از مرحله ۸ به مرحله ۲ و گرفتن داده جدید در پله زمانی بعدی ادامه می‌یابد. فلوچارت الگوریتم شناسایی برخط در شکل (۳-۲) ارائه شده است. تمام مراحل این الگوریتم غیرتکراری می‌باشند.



شکل (۳-۲) - دیاگرام بلوکی الگوریتم شناسائی برخط مدل‌های eTS

نام معیار	مقایسه اطلاعات داده جدید با مراکز موجود	بررسی فاصله
A	$p_k(z_k) > p_{\max}$	$\frac{p_k(z_k)}{p_{\max}} - \frac{\delta_{\min}}{\ \sigma_j\ } \geq 1$
B	$p_k(z_k) > p_{\max}$ or $0.5p_{\max} < p_k(z_k) < 0.675p_{\max}$	$\frac{\delta_{\min}}{\ \sigma_i\ } < 0.5$
C	$p_k(z_k) > p_{\max}$ or $p_k(z_k) < p_{\min}$	$\frac{\delta_{\min}}{\ \sigma_i\ } < 0.5$
D	$p_k(z_k) > p_{\max}$ or $p_k(z_k) < p_{\min}$	$\frac{\delta_{\min}}{\ \sigma_i\ } < 0.5$ or $\frac{\delta_{\min}}{\ \sigma_i\ } > 0.85$
E	$p_k(z_k) > p_{\max}$	$\frac{p_k(c_j)}{p_{\max}} + \frac{\delta_{\min}}{\ \sigma_j\ } < 1$
F	$\varphi_j \leq F_{gen}$	$\frac{\delta_{\min}}{\ \sigma_i\ } < 0.5$
G	$p_k(z_k) > p_{\max}$	$\frac{\delta_{\min}}{\ \sigma_i\ } < 0.5$

جدول ۳-۱- معیارهای مورد استفاده برای اصلاح پایگاه قواعد

جدول (۳-۱) شرایط معیارهای مورد استفاده برای اصلاح پایگاه قواعد را بیان می‌کند. معیارهای ستون اول برای سنجش میزان اطلاعات (پتانسیل اطلاعاتی) نقطه جدید و مقایسه آن با پتانسیل اطلاعاتی مراکز خوشه‌های موجود بکار می‌روند. معیارهای ستون دوم نیز میزان فاصله نقطه جدید تا مراکز خوشه‌های قبلی را بررسی می‌کنند. در این جدول داریم:

$$p_{\min} = \min_{i=1}^R p_k(c_i) \quad (70-3)$$

در معیار F بجای پتانسیل از سطوح فعالسازی قواعد و تعداد آنها بعنوان معیار سنجش کیفیت آنها استفاده می‌شود. با وارد شدن هر داده جدید، قاعده دارای بیشترین سطح فعالسازی یافته می‌شود:

$$\mathcal{J} = \arg \max_{i=1}^R (\varphi_i) \quad (۷۱-۳)$$

اگر $\varphi_j \leq F_{gen}$ ، آنگاه قاعده جدیدی ایجاد می‌شود . با توجه به تغییر تعداد قواعد در حین اجرای الگوریتم ، F_{gen} ثابت نیست و بصورت خودکار متناسب با تعداد قواعد تنظیم می‌شود . در ابتدای الگوریتم بعلت کوچک بودن R ، F_{gen} بزرگ می‌باشد تا شرط ایجاد قاعده جدیدی براحتی برقرار شده و تعداد قواعد افزایش یابد. همچنین با توجه به بزرگ بودن F_{gen} ، مقادیر بزرگی برای عرض توابع عضویت بدست می‌آید. بدین صورت یک توصیف کلی و تقریبی از الگوی توزیع داده‌ها بدست می‌آید . سپس با افزایش تعداد قواعد پارامتر F_{gen} کاهش می‌یابد تا از افزایش بی‌رویه تعداد قواعد جلوگیری شود. عرض توابع عضویت نیز کاهش می‌یابد .

در روشهای شناسائی برخط معمولی (نظیر روش RLS) باید سیگنال ورودی تا حد امکان همه فرکانسهای طبیعی سیستم را تحریک کند . عبارت دیگر در حین آزمایش اگر فرکانس جدیدی از سیستم تحریک شود که در داده‌های آموزشی تحریک نشده باشد روشهای متداول در پیش‌بینی رفتار آینده سیستم خطای قابل توجهی خواهند داشت. این مشکل بیشتر بخاطر ثابت بودن ساختار مدل‌های متداول است که سعی می‌کنند تنها با تغییر دادن پارامترها در یک ساختار ثابت ، خود را بر سیستم واقعی منطبق کنند. بعلت اینکه پارامترهای مدل‌های متداول بصورت بازگشتی محاسبه می‌شوند بنابراین ایجاد تغییر ناگهانی در این پارامترها (جهت انطباق با شرایط کاملاً جدید) امکانپذیر نیست. گذشته از این معمولاً از سیگنال PRBS برای تحریک استفاده می‌شود که یک سیگنال پله‌ای (با عرض و مدت زمان پله‌های تصادفی) است و تحریک سیستم‌های عملی با این سیگنال مشکل ، هزینه‌بر و گاهی خطرناک می‌باشد.

یکی از مزایای مدلسازی با الگوریتم ارائه شده اینست که بر خلاف روشهای شناسائی سیستم متداول لزومی ندارد که سیگنال ورودی همه فرکانسهای طبیعی سیستم را تحریک کند چون در مدل TS بدست آمده نه تنها پارامترهای مدل بلکه ساختار آن نیز بصورت لحظه‌ای در حال اصلاح یافتن می‌باشد بنابراین در صورتی که تغییر پارامترها به تنهایی نتواند مدل سیستم را بر شرایط ایجاد شده جدید (ناشی از تحریک فرکانس جدید) منطبق کند آنگاه ساختار مدل اصلاح شده و قاعده جدیدی اضافه می‌شود که مخصوصاً جهت نمایش فرکانس جدید به مدل افزوده شده است.

۳-۸- شرایط تضمین پایداری الگوریتم کاهش گرادیان

می‌دانیم که الگوریتمهای متداول روشهای شناسائی فازی-عصبی (مثلاً کاهش گرادیان و کمترین مربعات) در شرایط ایده‌آل پایدار هستند. اما وجود دینامیک مدل نشده ممکن است باعث رانش پارامترها و ناپایداری این الگوریتمها شود.

در این بخش قوانین یادگیری جدیدی برای شبکه‌های فازی-عصبی نوع ممدانی و TS ارائه می‌شوند. در این قوانین، سرعت‌های یادگیری متغیر با زمان بوده و از روی داده‌های ورودی-خروجی و ساختار مدل تعیین می‌شوند. برای محاسبه سرعت‌های یادگیری هیچگونه نیازی به اطلاعات اولیه از جمله محدوده‌های خطای تخمین وجود ندارد. مدل فازی ممدانی نوعی بصورت زیر بیان می‌شود:

$$R^i: IF \ x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \text{ and } \dots \ x_n \text{ is } A_{in} \quad (72-3)$$

$$THEN \ \hat{y}_1 \text{ is } B_{i1} \text{ and } \dots \text{ and } \hat{y}_m \text{ is } B_{im}$$

برای انجام نگاشت از بردار ورودی $X = [x_1 \dots x_n] \in \mathfrak{R}^n$ به بردار خروجی $\hat{Y}(k) = [\hat{y}_1 \dots \hat{y}_m]^T \in \mathfrak{R}^{m \times 1}$ از قاعده فازی استفاده می‌کنیم. A_{i1}, \dots, A_{in} و B_{i1}, \dots, B_{im} بترتیب مجموعه‌های فازی استاندارد ورودی و خروجی می‌باشند. هر متغیر ورودی دارای R_i مجموعه فازی است که در حالت کامل بودن اتصالات خواهیم داشت:

$$R = R_1 \times R_2 \times \dots \times R_n \quad (73-3)$$

با استفاده از سیستم استنتاج حاصلضرب، فازی‌ساز منفرد و غیرفازی‌ساز میانگین مراکز خروجی p ام سیستم فازی بصورت زیر بیان می‌شود:

$$\hat{y}_p = \frac{\left(\sum_{i=1}^R w_{ip} \left[\prod_{j=1}^n \mu_{A_{ij}} \right] \right)}{\left(\sum_{i=1}^R \left[\prod_{j=1}^n \mu_{A_{ij}} \right] \right)} \quad (74-3)$$

در رابطه بالا w_{ip} نقطه‌ای است که در آن $\mu_{B_i} = 1$ با استفاده از تعریف آستانه تحریک نرمالیزه شده عبارت (۷۴-۳) بصورت زیر بیان خواهد شد:

$$\hat{Y}(k) = W(k)\Phi(X(k)) \quad (75-3)$$

در رابطه بالا داریم:

$$W(k) = \begin{bmatrix} W_{11} & & W_{R1} \\ & \ddots & \\ W_{1m} & & W_{Rm} \end{bmatrix} \in R^{m \times 1} \quad (76-3)$$

$$\Phi[X(k)] = [\lambda_1 \dots \lambda_R]^T \in R^{R \times 1}$$

در مدل فازی TS نیز خروجی q ام را می‌توان بفرم معادله (۷۵-۳) نوشت با این تفاوت که در اینجا داریم :

$$W(k) = \begin{bmatrix} P_{10}^1 \dots P_{10}^R & P_{11}^1 \dots P_{11}^R & \dots & P_{1n}^1 \dots P_{1n}^R \\ \vdots & \vdots & & \vdots \\ P_{m0}^1 \dots P_{m0}^R & P_{m1}^1 \dots P_{m1}^R & \dots & P_{mn}^1 \dots P_{mn}^R \end{bmatrix} \quad (77-3)$$

$$\Phi[X(k)] = [\lambda_1 \dots \lambda_R \quad x_1 \lambda_1 \dots x_1 \lambda_R \quad \dots \quad x_n \lambda_1 \dots x_n \lambda_R]^T \quad (78-3)$$

۳-۸-۱- بررسی پایداری در حالت معلوم بودن پارامترهای مقدمه

در این زیربخش فرض می‌کنیم که توابع عضویت مقدمه‌ها (A_{i1}, \dots, A_{in}) با استفاده از دانش اولیه معلوم هستند و در نتیجه λ_i ها قابل محاسبه می‌باشند. در این حالت مدل‌های ممدانی و TS بعلت معلوم بودن $\Phi[X(k)]$ بفرم یکسانی درمی‌آیند و تنها تفاوت آنها در تعریف $W(k)$ می‌باشد. خطای شناسائی را بصورت زیر تعریف می‌کنیم:

$$e(k) = \hat{Y}(k) - Y(k) \quad (79-3)$$

مطابق تئوری تقریب توابع ، منطق فازی و شبکه‌های عصبی، فرایند غیرخطی شناسائی شده را می‌توان بصورت زیر نمایش داد :

$$Y(k) = \mathcal{W}^* \Phi[X(k)] - \mu(k) \quad (80-3)$$

\mathcal{W}^* ضرایب نامعلومی هستند که می‌توانند دینامیک مدل نشده $\mu(k)$ را به حداقل برسانند. با استفاده از (۷۹-۳) و (۸۰-۳) ، خطای شناسائی بصورت زیر بیان می‌شود :

$$e(k) = \tilde{W}(k) \Phi[X(k)] + \mu(k) \quad (81-3)$$

$$\tilde{W}(k) = W(k) - \mathcal{W}^* \quad (82-3)$$

در این بخش تنها به شناسائی حلقه‌باز پرداخته می‌شود و فرض می‌کنیم که سیستم واقعی دارای پایداری BIBO باشد. بعلت محدود بودن ورودی، خروجی و Φ ، $\mu(k)$ نیز محدود خواهد بود. قضیه زیر الگوریتم کاهش گرادیان پایدار برای مدلسازی فازی-عصبی را بیان می‌کند:

قضیه ۱: اگر از سیستم فازی-عصبی (۳-۷۴) برای شناسائی سیستم غیرخطی استفاده کنیم آنگاه الگوریتم کاهش گرادیان زیر با استفاده از سرعت یادگیری متغیر با زمان باعث محدود شدن خطای شناسائی می‌شود:

$$W(k+1) = W(k) - \eta_k e(k) \Phi^T [X(k)] \quad (۳-۸۳)$$

که اسکالر μ_k (سرعت یادگیری متغیر با زمان) از رابطه زیر بدست می‌آید:

$$\mu_k = \left(\eta + \|\Phi[X(k)]\|^2 \right)^{-1}; \quad 0 < \eta < 1 \quad (۳-۸۴)$$

خطای شناسائی نرمالیزه شده که بصورت زیر تعریف می‌شود:

$$e_N(k) = \frac{e(k)}{\left(1 + \max_k \|\Phi[X(k)]\|^2 \right)} \quad (۳-۸۵)$$

می‌تواند در رابطه عملکرد میانگین زیر صدق کند:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{k=1}^T \|e_N(k)\|^2 \leq \bar{\mu} \quad (۳-۸۶)$$

$$\bar{\mu} = \max_k \left[\|\mu(k)\|^2 \right] \quad \text{که}$$

۳-۸-۲- بررسی پایداری در حالت یادگیری پارامترهای مقدمه

در حالت کلی قبل از انجام شناسائی هیچکدام از پارامترهای مقدمه و نتیجه مشخص نیستند. خروجی

q ام شبکه فازی-عصبی را می‌توان بصورت زیر نوشت:

$$\hat{y}_q = \frac{\sum_{i=1}^R w_{iq} \left[\prod_{j=1}^n \exp \left(-\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2} \right) \right]}{\sum_{i=1}^R \left[\prod_{j=1}^n \exp \left(-\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2} \right) \right]} \quad (۳-۸۷)$$

اجازه دهید فرض کنیم :

$$\varphi_i = \prod_{j=1}^n \exp\left(-\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}\right) ; a_q = \sum_{i=1}^R w_{iq} \varphi_i ; b = \sum_{i=1}^R \varphi_i \quad (۸۸-۳)$$

بنابراین خواهیم داشت :

$$\hat{y}_q = \frac{a_q}{b} \quad (۸۹-۳)$$

مشابه رابطه (۷۴-۳) ، فرایند غیرخطی را می‌توان بصورت زیر بیان کرد :

$$\hat{y}_q = \frac{\sum_{i=1}^R w_{iq}^* \left[\prod_{j=1}^n \exp\left(-\frac{(x_j - c_{ij}^*)^2}{\sigma_{ij}^{*2}}\right) \right]}{\sum_{i=1}^R \left[\prod_{j=1}^n \exp\left(-\frac{(x_j - c_{ij}^*)^2}{\sigma_{ij}^{*2}}\right) \right]} \quad (۹۰-۳)$$

c_{ij}^* ، σ_{ij}^{*2} و w_{iq}^* پارامترهای نامعلومی هستند که باید دینامیک مدل نشده را به حداقل برسانند.

اگر f تابع همواری از سه متغیر باشد آنگاه بسط سری تیلور آن بصورت زیر نوشته خواهد شد:

$$f(x_1, x_2, x_3) = \sum_{k=0}^{R-1} \frac{1}{k!} \left[(x_1 - x_1^0) \frac{\partial}{\partial x_1} + (x_2 - x_2^0) \frac{\partial}{\partial x_2} + (x_3 - x_3^0) \frac{\partial}{\partial x_3} \right]^k f + M_R \quad (۹۱-۳)$$

در این رابطه M_R باقیمانده سری تیلور می‌باشد. اگر بجای x_1 ، x_2 و x_3 بترتیب از w_{ip}^* ، c_{ip}^*

و σ_{ij}^{*2} استفاده کنیم آنگاه x_1^* ، x_2^* و x_3^* متناظر با w_{ip}^* ، c_{ip}^* و σ_{ij}^{*2} خواهد بود و داریم :

$$y_q + \mu_q = \hat{y}_q + \sum_{i=1}^l \frac{(w_{iq}^* - w_{iq}) z_i}{b} + \sum_{i=1}^R \sum_{j=1}^n \frac{\partial}{\partial c_{ij}} \left(\frac{a_q}{b} \right) (c_{ij}^* - c_{ij}) + \sum_{i=1}^R \sum_{j=1}^n \frac{\partial}{\partial \sigma_{ij}} \left(\frac{a_q}{b} \right) (\sigma_{ij}^* - \sigma_{ij}) + M_{1q} \quad (۹۲-۳)$$

که M_{1q} خطای تقریب مرتبه اول سری تیلور می‌باشد ($q = 1, \dots, m$). با استفاده از قاعده زنجیره‌ای

خواهیم داشت :

$$\begin{aligned} \frac{\partial}{\partial c_{ij}} \left(\frac{a_q}{b} \right) &= \frac{\partial}{\partial z_i} \left(\frac{a_q}{b} \right) \frac{\partial z_i}{\partial c_{ij}} \\ &= \left(\frac{1}{b} \frac{\partial a_q}{\partial z_i} + \frac{\partial}{\partial z_i} \left(\frac{1}{b} \right) a_q \right) \left(2z_i \frac{x_j - c_{ij}}{\sigma_{ij}^2} \right) \\ &= \left(\frac{w_{iq} - a_q}{b} \right) \left(2z_i \frac{x_j - c_{ij}}{\sigma_{ij}^2} \right) = 2\varphi_i \frac{w_{iq} - \hat{y}_q}{b} \frac{x_j - c_{ij}}{\sigma_{ij}^2} \end{aligned} \quad (93-3)$$

$$\frac{\partial}{\partial \sigma_{ij}} \left(\frac{a_q}{b} \right) = \frac{\partial}{\partial \varphi_i} \left(\frac{a_q}{b} \right) \frac{\partial z_i}{\partial \sigma_{ij}} = 2\varphi_i \frac{w_{iq} - \hat{y}_q}{b} \frac{(x_j - c_{ij})^2}{\sigma_{ij}^3} \quad (94-3)$$

بیان ماتریسی روابط بالا بصورت زیر خواهد بود :

$$y_q + \mu_q = \hat{y}_q - Z(k) \tilde{w}_q - D_{z_q} \bar{C}_k E - D_{z_q} \bar{B}_k E - M_{1q} \quad (95-3)$$

در رابطه بالا داریم :

$$Z(k) = \left[\frac{\varphi_1}{b} \dots \frac{\varphi_R}{b} \right]^T ; \quad w_q = [w_{1q} \dots w_{Rq}] ; \quad \tilde{w}_q = w_q - w_q^*$$

$$D_{z_q} = \left[2z_1 \frac{w_{1q} - \hat{y}_q}{b}, \dots, 2z_R \frac{w_{Rq} - \hat{y}_q}{b} \right] ; \quad E = [1, \dots, 1]^T$$

$$\bar{C}_k = \begin{bmatrix} \frac{x_1 - c_{11}}{\sigma_{11}^2} (c_{11} - c_{11}^*) & \dots & \frac{x_n - c_{1n}}{\sigma_{1n}^2} (c_{1n} - c_{1n}^*) \\ \dots & \dots & \dots \\ \frac{x_1 - c_{1R}}{\sigma_{1R}^2} (c_{R1} - c_{R1}^*) & \dots & \frac{x_n - c_{nR}}{\sigma_{nR}^2} (c_{Rn} - c_{Rn}^*) \end{bmatrix}$$

$$\bar{B}_k = \begin{bmatrix} \frac{(x_1 - c_{11})^2}{\sigma_{11}^3} (\sigma_{11} - \sigma_{11}^*) & \dots & \frac{(x_n - c_{1n})^2}{\sigma_{1n}^3} (\sigma_{1n} - \sigma_{1n}^*) \\ \dots & \dots & \dots \\ \frac{(x_1 - c_{R1})^2}{\sigma_{R1}^3} (\sigma_{R1} - \sigma_{R1}^*) & \dots & \frac{(x_n - c_{Rn})^2}{\sigma_{Rn}^3} (\sigma_{Rn} - \sigma_{Rn}^*) \end{bmatrix}$$

با تعریف خطای شناسائی بصورت $e_q = \hat{y}_q - y_q$ و استفاده از رابطه (95-3) خواهیم داشت :

$$e_q = Z(k) \tilde{w}_k + D_{z_q} \bar{C}_k E + D_{z_q} \bar{B}_k E + \mu_q - M_{1q} \quad (96-3)$$

بیان برداری رابطه بالا را بصورت داریم :

$$e(k) = Z(k) \tilde{w}_k + D_z \bar{C}_k E + D_z \bar{B}_k E + \xi(k) \quad (97-3)$$

که در این رابطه تعاریف زیر استفاده شده‌اند :

$$e(k) = [e_1 \cdots e_m]^T$$

$$\tilde{w}_k = \begin{bmatrix} w_{11} - w_{11}^* & w_{1m} - w_{1m}^* \\ w_{R1} - w_{R1}^* & w_{Rm} - w_{Rm}^* \end{bmatrix}$$

$$D_z = \begin{bmatrix} \frac{x_1 - c_{11}}{\sigma_{11}^2} (c_{11} - c_{11}^*) & \frac{x_n - c_{1n}}{\sigma_{1n}^2} (c_{1n} - c_{1n}^*) \\ \vdots & \vdots \\ \frac{x_1 - c_{1R}}{\sigma_{1R}^2} (c_{R1} - c_{R1}^*) & \frac{x_n - c_{nR}}{\sigma_{nR}^2} (c_{Rn} - c_{Rn}^*) \end{bmatrix}$$

$$\xi(k) = \mu - M_1 \quad ; \quad \mu = [\mu_1 \cdots \mu_m]^T \quad ; \quad M_1 = [M_{11} \cdots M_{1m}]^T$$

با توجه به محدود بودن تابع گوسی φ و پایداری BIBO فرایند μ و M_1 محدود خواهند بود. بنابراین $\xi(k)$ در رابطه (97-3) نیز پایدار خواهد بود. قضایای زیر شرایط پایداری الگوریتم‌های آموزش شبکه‌های فازی-عصبی زمان‌گسسته نوع ممدانی و TS ارائه می‌کنند:

قضیه ۲: اگر از شبکه فازی-عصبی نوع ممدانی یا TS برای شناسائی سیستم غیرخطی استفاده شود الگوریتم پس‌انتشار زیر خطای شناسائی را محدود خواهد کرد :

$$c_{ij}(k+1) = c_{ij}(k) - 2\eta_k \varphi_i \frac{w_{ip} - \hat{y}_p}{b} \frac{x_j - c_{ij}}{\sigma_{ij}^2} (\hat{y}_q - y_q) \quad (98-3)$$

$$\sigma_{ij}(k+1) = \sigma_{ij}(k) - 2\eta_k \varphi_i \frac{w_{ip} - \hat{y}_p}{b} \frac{(x_j - c_{ij})^2}{\sigma_{ij}^3} (\hat{y}_q - y_q) \quad (99-3)$$

که در رابطه بالا داریم :

$$\eta_k = \eta + \|Z\|^2 + 2\|D_z\|^2 \quad ; \quad 0 < \eta < 1$$

میانگین خطای شناسائی در رابطه زیر صدق می‌کند :

$$\mathcal{J} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{k=1}^T \|e(k)\|^2 \leq \frac{\eta}{\pi} \bar{\xi} \quad (۱۰۰-۳)$$

در رابطه () تعاریف زیر را خواهیم داشت :

$$\pi = \frac{\eta}{(1+\kappa)^2} > 0 \quad ; \quad \kappa = \max_k \left(\|Z\|^2 + 2\|D_z\|^2 \right) \quad ; \quad \bar{\xi} = \max_k \left[\|\xi(k)\|^2 \right]$$

قضایای ۱ و ۲ در مرجع [۴۶] اثبات شده‌اند. توجه شود که برای تضمین پایداری یادگیری، شرط $0 < \eta < 1$ یک شرط لازم ولی ناکافی می‌باشد. امکان بالابردن سرعت الگوریتم با استفاده از $\eta > 1$ وجود دارد اما دیگر نمی‌توان از پایدار ماندن الگوریتم در تمام مدت شناسائی اطمینان کامل داشت.

۳-۹- شبیه‌سازی

۳-۹-۱- پیش‌بینی سری زمانی مکی-گلاس^۱

داده‌های ورودی- خروجی سری زمانی آشوبناک^۲ مکی-گلاس در بسیاری از مراجع شناسائی سیستم بعنوان معیاری برای بررسی عملکرد الگوریتمهای شناسائی بکار رفته‌اند. این سری بوسیله معادله دیفرانسیل تاخیری مکی-گلاس زیر تولید می‌شود :

$$\dot{x}(t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t).$$

الگوریتم ارائه شده روی مساله پیش‌بینی سری زمانی مکی-گلاس مورد آزمایش قرار گرفته است. مقادیر $x(0)$ و τ بترتیب برابر $1/2$ و 17 در نظر گرفته می‌شوند و مقدار $x(t+85)$ بر اساس مقادیر سیگنال x در لحظه جاری و لحظات مربوط به $6, 12, 18$ پله زمانی عقب‌تر تخمین زده می‌شود بنابراین ورودیها و خروجیها بصورت زیر خواهند بود :

$$u(t) = [x(t-18) \quad x(t-12) \quad x(t-6) \quad x(t)]$$

$$y(t) = x(t+85)$$

^۱ Mackey-glass
^۲ Chaotic

نتایج اعمال الگوریتم مدل‌سازی به‌همراه معیارهای مختلف جدول (۳-۱) در جداول زیر نمایش داده شده است. پارامتر NDEI بصورت نسبت RMSE به انحراف استاندارد داده‌های اصلی تعریف می‌شود.

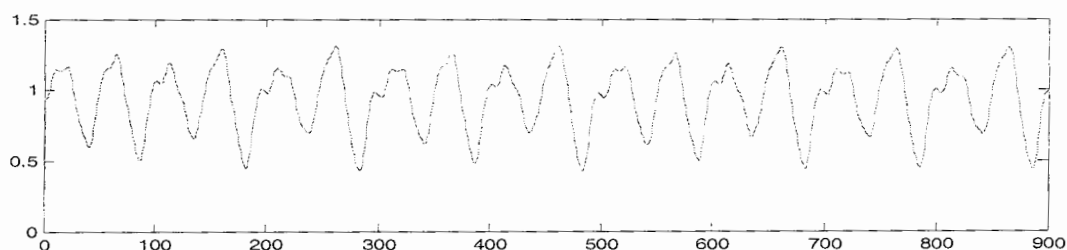
معیار	A	B	C	D	E	F	G
RMSE	۰/۰۰۲۲	۰/۰۰۲۷	۰/۰۰۲۳	۰/۰۳۰۵	۰/۰۰۲۲	۰/۰۳۱	۰/۰۰۲۲
NDEI	۰/۰۰۹۸	۰/۰۱۱۷	۰/۰۱۰۳	۰/۱۳۴۴	۰/۰۰۹۸	۰/۰۱۳۵	۰/۰۰۹۸
تعداد قواعد	۱۰	۱۷	۵	۱	۱۰	۵۹	۳

جدول ۲-۲ - عملکرد مدل eTS در پیش‌بینی سری مکی-گلاس (حالت D ثابت)

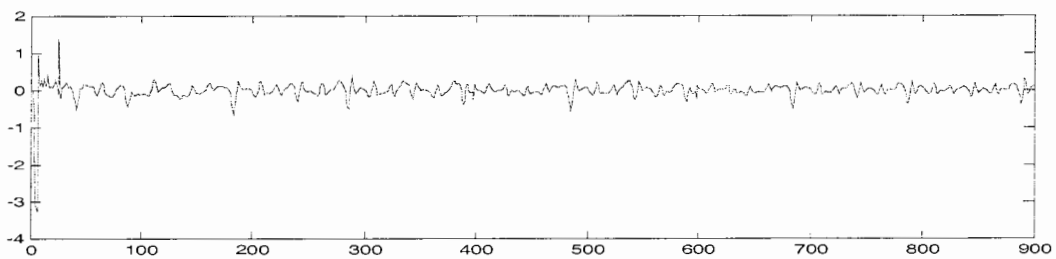
معیار	A	B	C	D	E	F	G
RMSE	۰/۰۰۲۱	۰/۰۳۰۵	۰/۰۳۰۵	۰/۰۳۰۵	۰/۰۰۲۱	۰/۰۰۵۵	۰/۰۰۲۶
NDEI	۰/۰۰۹۱	۰/۱۳۴۴	۰/۱۳۴۴	۰/۱۳۴۴	۰/۰۰۹۱	۰/۰۲۴۱	۰/۰۱۱۴
تعداد قواعد	۵	۲	۲	۱	۵	۲	۲

جدول ۳-۳ - عملکرد مدل eTS در پیش‌بینی سری زمانی مکی-گلاس (حالت D متغیر)

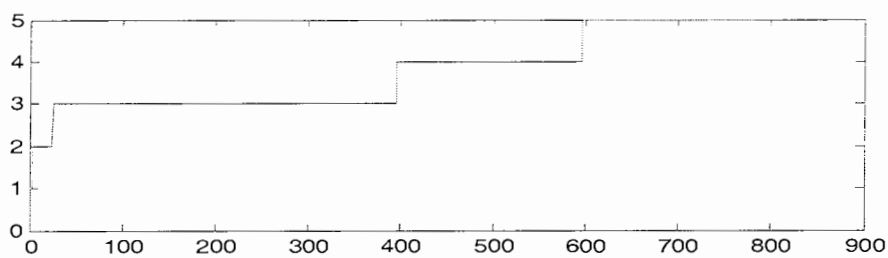
در جدول (۳-۴) نتایج بدست آمده از شناسائی سری مکی-گلاس با نتایج حاصل از روشهای دیگر آموزش مدل‌های TS (برگرفته از مراجع ۳۱ و ۳۴) مقایسه شده‌اند. ملاحظه می‌شود که تعداد قواعد مدل فازی بدست آمده از تمام مدل‌های مراجع مذکور کمتر است که این فشردگی مدل در پیاده‌سازی سخت‌افزاری و کاهش حجم محاسبات بسیار مفید است در عین حال دقت مدل نیز به مراتب از دقت دیگر مدل‌ها بیشتر است.



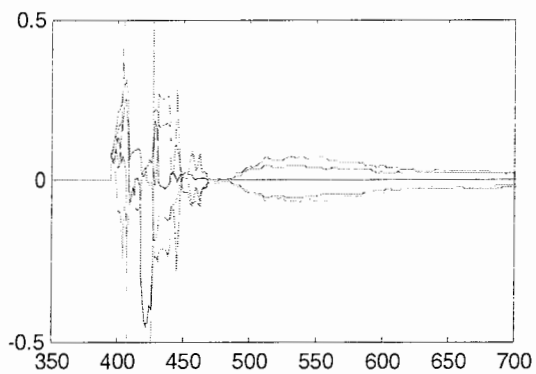
شکل ۳-۳ - پیش‌بینی سری زمانی مکی-گلاس (داده اصلی —، مقادیر پیش‌بینی شده ---)



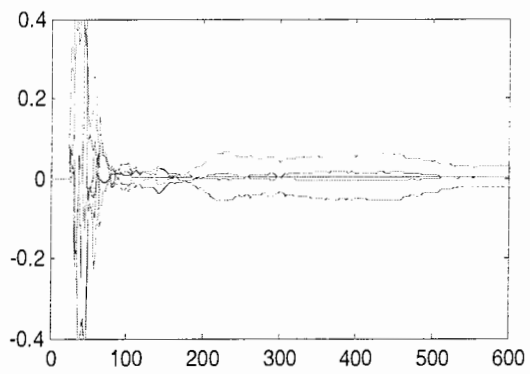
شکل ۳-۴- درصد خطا در پیش‌بینی سری زمانی مکی-گلاس با استفاده از مدل فازی TS



شکل ۳-۵- تعداد قواعد مدل فازی TS



شکل ۳-۷- پارامترهای قسمت نتیجه قاعده چهارم



شکل ۳-۶- پارامترهای قسمت نتیجه قاعده سوم

Algorithm/authors	Rules/ units	NDEI
RAN(kasabov&song , ۲۰۰۲)	۱۱۳	۰/۰۳۷۳
ESOM(kasabov&song,۲۰۰۲)	۱۱۴	۰/۳۲
ESOM(kasabov&song,۲۰۰۲)	۱۰۰۰	۰/۰۴۴
EfuNN (kasabov&song,۲۰۰۲)	۱۹۳	۰/۳۰۱
EfuNN (kasabov&song,۲۰۰۲)	۱۱۲۵	۰/۰۹۰۴
Denfis (kasabov&song,۲۰۰۲)	۸۸۳	۰/۰۳۳
Denfis (kasabov&song,۲۰۰۲)	۵۸	۰/۲۷۶
eTS[۱]	۱۱۳	۰/۰۹۵
eTS (victor&durado , ۲۰۰۳) [۳]	۹	۰/۳۸
Fuzzy Transversal filter[۴]	-	۰/۰۵۹۷
SONFIN [c.juang & c.lin , ۱۹۹۸]	۹	۰/۰۷۹۶
Neural gas	۱۰۰۰	۰/۰۶۲
DFNN [۵]	۵	۰/۰۵۸۴
OLS [۵]	۱۳	۰/۰۶۹۸
RBF AFS[۵]	۱۶	۰/۰۴۷۳
مدل فازی شناسائی شده با الگوریتم بخش ۴-۷	۴	۰/۰۱۱۴

جدول ۳-۴- مقایسه عملکرد الگوریتم ارائه شده با نتایج برخی مقالات

۳-۹-۲- پیش‌بینی داده‌های کوره گازی Box-Jenkins

مجموعه داده Box-Jenkins یکی از معروفترین مجموعه‌های معیار برای آزمایش دقت روشهای شناسائی می‌باشد و شامل ۲۹۰ جفت داده ورودی-خروجی است که با پیروی نمونه‌برداری ۹ S از یک کوره آزمایشگاهی جمع‌آوری شده‌اند. نرخ جریان متان بعنوان متغیر ورودی و غلظت CO_2 متغیر خروجی می‌باشد.

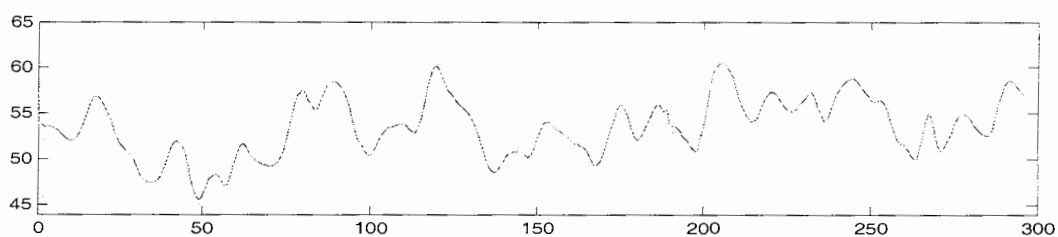
در این قسمت الگوریتم شناسایی ساده‌سازی شده را به این مجموعه داده اعمال می‌کنیم و یک مدل فازی تکاملی برای این داده‌ها بدست می‌آوریم. شکل (۴-۸) داده‌های اصلی و پیش‌بینی انجام شده توسط مدل فازی (معیار E) را نشان می‌دهد.

معیار	A	B	C	D	E	F	G
RMSE	۰/۳۴۷۷	۰/۳۵۷۵	۰/۳۴۷۹	۱/۰۱۲۹	۰/۳۴۷۷	۰/۴۲۳۱	۰/۳۵۶۳
NDEI	۰/۱۰۸۵	۰/۱۱۶۶	۰/۱۰۸۶	۰/۳۱۶۲	۰/۱۰۸۵	۰/۱۳۲۱	۰/۱۱۱۲
تعداد قواعد	۱۳	۳	۴	۱	۱۳	۳	۲

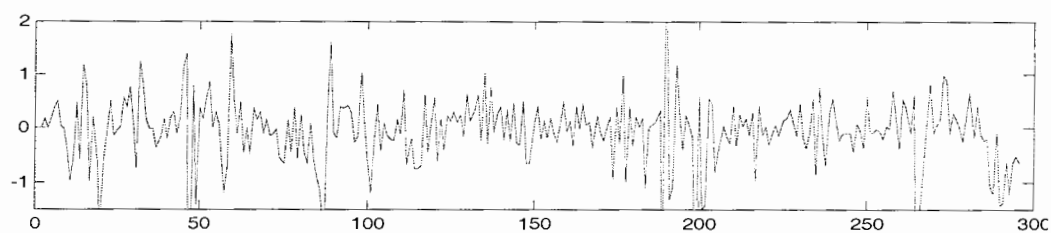
جدول ۳-۵ - عملکرد مدل eTS در پیش‌بینی داده‌های کوره گازی (حالت D ثابت)

معیار	A	B	C	D	E	F	G
RMSE	۰/۳۷۶۹	۱/۰۱۲۹	۱/۰۱۲۹	۱/۰۱۲۹	۰/۳۷۶۹	۰/۳۴۵۸	۰/۳۵۶۳
NDEI	۰/۱۱۷۷	۰/۳۱۶۲	۰/۳۱۶۲	۰/۳۱۶۲	۰/۱۱۷۷	۰/۱۰۷۹	۰/۱۱۱۲
تعداد قواعد	۵	۲	۲	۱	۵	۲	۲

جدول ۳-۶ - عملکرد مدل eTS در پیش‌بینی داده‌های کوره گازی (حالت D متغیر)



شکل ۳-۸ - شناسایی کوره گازی Box-Jenkins (داده اصلی —، مقادیر پیش‌بینی شده ---)



شکل ۳-۹ - درصد خطا در شناسایی کوره گازی Box-Jenkins با استفاده از مدل فازی TS

۳-۹-۳- شناسائی سیستم CSTR

در این قسمت الگوریتم مدلسازی ارائه شده را به سیستم CSTR اعمال کرده و مدلی بدست می-آوریم که با استفاده از داده‌های I/O اندازه‌گیری شده می‌تواند خروجی سیستم را در یک افق پیش-بینی H_2 تخمین بزند :

$$\hat{y}(t+H_2) = f(y(t), \dots, y(t-m), u(t-d), \dots, u(t-d-n)) \quad (10-1-3)$$

هدف ما پیش‌بینی خروجی در ۸۰ پله زمانی جلوتر با استفاده از مقادیر ورودی و خروجی تا لحظه حاضر می‌باشد ($H_{2p} = 80$). مقادیر مناسب برای m و n با سعی و خطا برابر ۳ و ۱۰ بدست آمده‌اند. هر دو ورودی سیستم CSTR بصورت سیگنالهای PRBS^۱ در نظر گرفته شده‌اند تا فرکانسهای سیستم بطرز مناسبی تحریک شوند البته همانطور که قبلاً نیز اشاره شد الگوریتم ارائه شده حتی در صورت تحریک مدهای تجربه نشده نیز بخوبی می‌تواند خروجی را پیش‌بینی کند و برای داشتن یک پیش‌بینی خوب در هر مرحله لازم نیست تمام فرکانسهای طبیعی سیستم از قبل بوسیله سیگنال ورودی تحریک شده باشند. در اینجا هدف از تحریک سیستم CSTR با سیگنالهای PRBS، مجبور کردن سیستم به نمایش همه خواص غیرخطی خود بوده است. نتایج اعمال معیارهای مختلف به داده‌های ورودی-خروجی جمع‌آوری شده از سیستم در جدولهای (۳-۷) و (۳-۸) نشان داده شده‌اند.

معیار	A	B	C	D	E	F	G
RMSE	۰/۱۴۲۴	۰/۶۸۷۶	۴/۸۳۷۶	۰/۰۱۸۳	۰/۱۴۲۴	۰/۰۲۴۶	۰/۰۲۶۹
NDEI	۶/۱۷۷۶	۲۶/۸۳۵۹	۲۰۹/۸۹۶	۰/۷۹۴۳	۶/۱۷۷۶	۱/۰۶۹۶	۱/۱۶۹۶
تعداد قواعد	۱۵	۴۸	۵۱	۱	۱۵	۱۳	۷

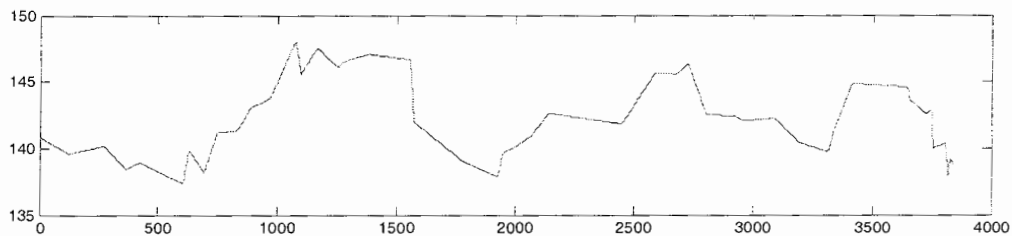
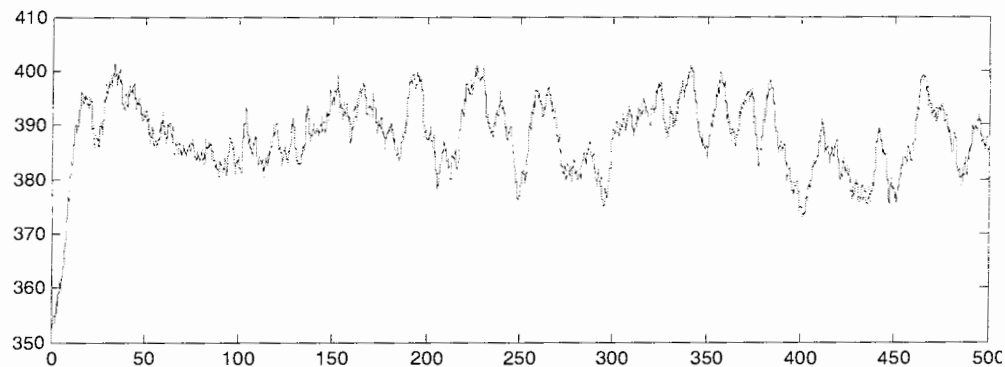
جدول ۳-۷- نتایج شناسائی سیستم CSTR (حالت D ثابت)

^۱ Probabilistic Random Binary Sequence

معیار	A	B	C	D	E	F	G
RMSE	۰/۰۹۶۱	۰/۱۵۸۳	۰/۱۱۴۳	۰/۰۱۸۳	۰/۱۱۴۸	۰/۰۲۵۸	۰/۰۲۴۹
NDEI	۴/۱۶۸۷	۶/۸۶۶۸	۴/۹۵۷۳	۰/۷۹۴۳	۴/۹۸۰۱	۱/۱۲۱۷	۱/۰۸۲۶
تعداد قواعد	۱۲	۱۴	۵	۱	۷	۴	۳

جدول ۳-۸- نتایج شناسایی سیستم CSTR (حالت D متغیر)

ملاحظه می‌شود که در هر دو حالت D متغیر و ثابت، معیار D به بهترین نتیجه دست یافته است. دقت معیارهای F و G نیز جالب توجه است اما معیارهای دیگر در مقایسه با سه معیار مذکور دقت خوبی نداشته‌اند. در مرحله بعد برای بررسی قابلیت الگوریتم ارائه شده در مدل‌سازی سیستم‌های متغیر با زمان فرض می‌کنیم که درجه حرارت ماده A در ورودی CSTR (T_0) دارای اغتشاش باشد (شکل ۳-۱۰). نتایج حاصل از اعمال الگوریتم مدل‌سازی با معیارهای D ، G و استفاده از D متغیر در جدول (۳-۹) نشان داده شده‌اند.

شکل ۳-۱۰- درجه حرارت ماده A در ورودی CSTR (T_0)شکل ۳-۱۱- خروجی CSTR (با ورودیهای PRBS و استفاده از T_0 متغیر)

وضعیت اغتشاش- معیار	RMSE) (بدون اغتشاش)	تعداد قواعد- بدون اغتشاش	RMSE) با (اغتشاش)	تعداد قواعد- با اغتشاش	درصد افزایش RMSE
D	۰/۰۱۸۳	۱	۰/۰۱۹۴	۲	٪ ۵/۶۷
G	۰/۰۲۱۹	۳	۰/۰۲۲۶	۵	٪ ۳/۰۹

جدول ۳-۹- نتایج شناسائی سیستم CSTR (حالت وجود اغتشاش در T_0)

نتایج فوق نشان می‌دهند که درصد افزایش RMSE در مورد هر دو معیار D و G از درصد تغییرات T_0 (با مقدار تقریبی ۷/۸٪) کمتر است که نشانه موفقیت الگوریتم ارائه شده در مدلسازی سیستمهای غیرخطی متغیر با زمان می‌باشد.

برای بررسی اثر معیار حذف قواعد بر عملکرد مدل ، الگوریتم شناسائی را بدون استفاده از معیار حذف تکرار می‌کنیم. مقایسه نتایج بدست آمده (جدول ۳-۱۰) با نتایج مندرج در جدول (۳-۸) نشان می‌دهد که در صورت عدم استفاده از معیار حذف، تعداد قواعد و در نتیجه حجم محاسبات بمقدار قابل توجهی افزایش می‌یابد در حالیکه افزایش دقت حاصله بسیار ناچیز است. بنابراین نتیجه می‌گیریم که حذف قواعد اضافی باعث فشردگی مدل و کاهش محاسبات شده و در عین حال دقت مدل محفوظ می‌ماند و کاهش حاصل شده در محاسبات پایگاه قواعد می‌تواند اثر افزایش محاسبات مربوط به اعمال معیار حذف قواعد را جبران کند.

معیار	A	B	C	D	E	F	G
RMSE	۰/۰۹۵۴	۰/۱۴۲۵	۰/۱۱۴۳	۰/۰۱۷۶	۰/۱۰۳۴	۰/۰۲۴۲	۰/۰۲۲۸
NDEI	۴/۱۱۲۳	۶/۱۲۴	۴/۳۲۶۲	۰/۷۲۲	۴/۲۰۴۵	۱/۱۱۹۲	۱/۰۷۳
تعداد قواعد	۱۷	۱۵	۸	۲	۸	۶	۴

جدول ۳-۱۰- نتایج شناسائی سیستم CSTR (بدون معیار حذف قواعد)

۳-۹-۴- نتایج بدست آمده از اعمال الگوریتم به مجموعه داده‌های مختلف

۱- معیارهای A و E در تمام موارد مشابه یکدیگر عمل می‌کنند (بغیر از شناسائی سیستم CSTR در حالت D متغیر).

۲- بیشتر معیارها در حالت ثابت بودن D به دقت پیش‌بینی بالاتری دست یافته‌اند اما در تمام موارد تعداد قواعد مدل نیز افزایش چشمگیری نشان می‌دهد. تعداد قواعد در برخی حالات ۴ تا ۸ برابر شده است و درصد افزایش دقت در تمام حالات از درصد افزایش تعداد قواعد کمتر بوده است. البته در مورد بیشتر معیارها، مدل فازی سیستم CSTR در حالت متغیر بودن D با تعداد قواعد کمتر به دقت بیشتری نسبت به حالت ثابت بودن D دست یافته است.

۳- معیارهای A و E در موارد زیادی بالاترین دقت را دارند اما تعداد قواعد آنها در حالت ثابت بودن D بیشتر از قواعد معیارهای دیگر است که این نکته نیز برتری حالت D متغیر را نشان می‌دهد.

۴- معیارهای B، C و D در مورد مجموعه‌های سری زمانی و کوره گازی با استفاده از D متغیر به نتایج مشابهی دست یافته‌اند اما از آنجا که در همین حالت معیارهای دیگر با تعداد قواعد مساوی دقت بالاتری داشته‌اند لذا توجه خود را به معیارهای دیگر معطوف می‌کنیم.

۵- در حالت D متغیر و در مورد هر سه مجموعه داده، معیارهای A و E نسبت به معیارهای G و F دارای تعداد قواعد بالاتری می‌باشند.

۶- با توجه به اینکه تحقق معیار اضافه شدن قواعد جدید (بخصوص در حالت D متغیر) پس از طی شدن حالات گذرا مشکل می‌شود لزومی ندارد که با وارد شدن هر داده جدید معیار حذف نیز اعمال شود بلکه می‌توانیم بصورت تناوبی بعد از وارد شدن هر N داده یکبار معیار مذکور را اعمال کنیم یا اینکه اعمال آنرا منوط به افزایش تعداد قواعد تا آستانه از پیش تعیین شده‌ای کنیم. با توجه به اینکه محاسبات پایگاه قواعد در هر لحظه انجام می‌شوند

در مجموع اگر بخواهیم مدل فازی هر دو ویژگی فشردگی ساختار (تعداد کم قواعد) و دقت بالا را با هم داشته باشد باید از D متغیر و یکی از معیارهای F یا G به‌مراه معیار حذف قواعد استفاده کنیم. در مورد سیستم CSTR می‌توان از معیار D نیز بخوبی استفاده کرد.

۳-۱۰- ساده‌سازی محاسباتی الگوریتم شناسائی مدل eTS

از آنجا که محاسبات روشهای شناسائی برخط در هر پریود نمونه‌برداری تکرار می‌شوند و در سیستمهای کنترل بلادرنگ غالباً پریود نمونه‌برداری کوچک است بنابراین کاهش حجم محاسبات الگوریتمهای شناسائی بسیار مطلوب می‌باشد. در این بخش چند راهکار جهت کاهش پیچیدگی محاسباتی فرایند شناسائی مدل eTS ارائه شده است تا پیاده‌سازی سخت‌افزاری این الگوریتم در سیستمهای کنترل بلادرنگ آسانتر و کم‌هزینه‌تر شود.

۱- بجای تابع عضویت گوسی از تابع معروف به کوشی^۱ استفاده می‌کنیم. این تابع تقریب سری تیلور مرتبه اول تابع گوسی بوده و محاسبه آن آسانتر است چون تنها شامل چهار عمل اصلی می‌باشد:

$$\mu_{ij}(x_j) = \left[1 + \left(\frac{2(x_j - c_{ij})}{\sigma_j} \right)^2 \right]^{-1} \quad (102-3)$$

همچنین می‌توانیم از توابع عضویت مثلثی نیز استفاده کنیم. در این حالت با فرض اینکه توابع عضویت گوسی هستند مرکز و شعاعهای آنها را بدست می‌آوریم. مراکز بدست آمده بدون تغییر بعنوان مراکز توابع عضویت مثلثی مورد استفاده قرار می‌گیرند و عرض هر تابع مثلثی بصورت ضربی از عرض تابع گوسی متناظر در نظر گرفته می‌شود.

۲- در خوشه‌سازی بجای پتانسیل اطلاعاتی، مفهوم پراکندگی^۲ را بکار می‌بریم. پراکندگی کلی^۳ هر نمونه بصورت میانگین فاصله بین آن نمونه و تمام نمونه‌های دیگر تعریف می‌شود:

$$S_k^G(z(k)) = \frac{\sum_{l=1}^N \sum_{j=1}^{n+m} (z_j(l) - z_j(k))^2}{N(n+m)} \quad (103-3)$$

مفهوم پراکندگی همانند پتانسیل شهودی است اما بعلاوه صحیح بودن اعداد موجود در مخرج رابطه پراکندگی (N و m, n)، پیاده‌سازی سخت‌افزاری تقسیم موجود بسادگی با استفاده از شیفت-رجیسترها انجام می‌شود حال آنکه در مورد پتانسیل مساله باین سادگی نیست زیرا اعداد موجود در

^۱ Cauchy

^۲ Scatter

^۳ Global

مخرج رابطه پتانسیل اعداد حقیقی هستند. پراکندگی محلی $z(k)$ (متعلق به خوشه i ام) تنها به فاصله بین آن داده تا داده‌های متعلق به خوشه i ام ($N_i(k)$) بستگی دارد:

$$S_k^I(z(k)) = \frac{\sum_{l=1}^{N_i(k)} \sum_{j=1}^{n+m} (z_j(l) - z_j(k))^2}{N_i(k)(n+m)} \quad (104-3)$$

جهت بدست آوردن رابطه بازگشتی پراکندگی کلی، انتقال ساده زیر را انجام می‌دهیم:

$$S_k^G(z(k)) = \frac{\sum_{l=1}^{k-1} \sum_{j=1}^{n+m} (z_j(l) - z_j(k))^2}{(k-1)(n+m)} \quad (105-3)$$

مشابه با حالت استفاده از پتانسیل اطلاعاتی، رابطه زیر برای پراکندگی اطلاعاتی بدست می‌آید:

$$S_k^G(z(k)) = \frac{(k-1) \sum_{j=1}^{n+m} z_j^2(k) - 2 \sum_{j=1}^{n+m} z_j(k) \beta_j(k) + \gamma(k)}{(k-1)(n+m)} \quad (106-3)$$

که در این رابطه داریم:

$$\gamma(k) = \sum_{l=1}^{k-1} \sum_{j=1}^{n+m} z_j^2(l) ; \beta_j(k) = \sum_{l=1}^{k-1} z_j(l) \quad (107-3)$$

اعداد موجود در مخرج عبارت (106-3) نیز همگی اعداد صحیح هستند حال آنکه مخرج عبارت متناظر برای محاسبه بازگشتی پتانسیل (رابطه 3-19) شامل اعداد حقیقی بود. روابط بازگشتی محاسبه $\gamma(k)$ و $\beta_j(k)$ از این قرار می‌باشند:

$$\begin{aligned} \gamma(k) &= \gamma(k-1) + \sum_{j=1}^{n+m} z_j^2 \\ \beta_j(k) &= \beta_j(k-1) + z_j(k-1) \end{aligned} \quad (108-3)$$

بعلت وابستگی پراکندگی هر داده به فاصله بین آن داده تا نقاط دیگر (شامل داده جدید)، با وارد شدن هر داده جدید باید پراکندگی مراکز خوشه‌ها بروزرسانی شود. بمنظور بدست آوردن یک رابطه بازگشتی برای پراکندگی نقطه کانونی c_i (مرکز خوشه i ام)، پراکندگی این نقطه در دو پله زمانی متوالی k و $k-1$ از یکدیگر تفریق می‌شوند :

$$S_k(z(k)) = \frac{\sum_{l=1}^{k-1} \sum_{j=1}^{n+m} (z_j(l) - z_j(k))^2}{(k-1)(n+m)} \quad (109-3)$$

$$S_{k-1}(z(k)) = \frac{\sum_{l=1}^{k-2} \sum_{j=1}^{n+m} (z_j(l) - z_j(k))^2}{(k-2)(n+m)} \quad (110-3)$$

از رابطه (۱۰۵-۳) داریم :

$$\sum_{l=1}^{k-2} \sum_{j=1}^{n+m} (z_j(l) - z_j(k-1))^2 = (k-2)(n+m)S_{k-1}(c_i) \quad (111-3)$$

همچنین از رابطه (۱۰۶-۴) خواهیم داشت :

$$S_k(c_i) = \frac{\sum_{l=1}^{k-2} \sum_{j=1}^{n+m} (z_j(l) - z_j(k))^2}{(k-1)(n+m)} + \frac{\sum_{j=1}^{n+m} (z_j(k-1) - z_j(k))^2}{(k-1)(n+m)} \quad (112-3)$$

و در نهایت رابطه زیر برای بروزرسانی بازگشتی پراکندگی نقاط کانونی بدست می‌آید :

$$S_k(c_i) = \frac{k-2}{k-1} S_{k-1}(c_i) + \frac{\sum_{j=1}^{n+m} (z_j(k) - z_j(k-1))^2}{(k-1)(n+m)} \quad (113-3)$$

۳- بجای معیار (۳-۶۹) از روش ساده‌سازی جمعیتی برای کاهش حجم پایگاه قواعد استفاده می‌کنیم بدین ترتیب که اگر تعداد نقاط متعلق به خوشه‌ای از یک درصد تعداد کل نقاط کمتر شود آنگاه آن خوشه و قاعده متناظر با آن از پایگاه قواعد حذف می‌شود. با توجه به اینکه در الگوریتمهای شناسائی برخط تعداد داده‌ها بطور پیوسته در حال افزایش است ممکن است در ذخیره‌سازی

تعداد نقاط مشکلاتی پیش آید. در صورت بروز این مشکل می‌توان از معیار دیگری برای تشخیص قواعد اضافی استفاده کرد بدین‌شکل که بجای ثبت تعداد نقاط هر خوشه تعداد لحظاتی را ثبت کنیم که هیچ داده‌ای وارد این خوشه نشده است. در صورتی که این عدد برای خوشه‌ای از یک آستانه (از پیش تعیین شده) بیشتر شد مشخص می‌شود که خوشه مذکور (حداقل در این لحظات) دیگر اهمیت سابق را ندارد و می‌توانیم آنرا حذف کنیم.

برای انعطاف بیشتر در روش تشخیص خوشه‌های اضافی می‌توانیم آستانه مربوط به هر خوشه را با متناسب با فاصله مرکز خوشه مذکور تا مراکز خوشه‌های مطرح در لحظات جاری در نظر بگیریم بدین ترتیب تفاوت بین خوشه‌هایی که در کل اهمیت خود را از دست داده‌اند با خوشه‌هایی که موقتاً کم-اهمیت شده‌اند بیشتر آشکار می‌شود.

یک مرحله مهم در پیاده‌سازی الگوریتم ساده‌شده، استانداردسازی برخط داده‌ها می‌باشد. در روشهای شناسائی خارج‌خط بعلت اینکه تمامی داده‌ها در ابتدای الگوریتم در دسترس هستند بنابراین محاسبه میانگین و انحراف استاندارد داده‌ها کار آسانی است اما استفاده از روابط معمولی برای محاسبه ایندو پارامتر در روشهای برخط بدلیل داشتن بار محاسباتی زیاد مقرون بصرفه نیست و بجای آن از روابط بازگشتی برای محاسبه عنصر z ام میانگین $(\bar{z}_j(k))$ و واریانس $(\sigma_j^2(k))$ استفاده می‌کنیم:

$$\bar{z}_j(k) = \frac{k-1}{k} \bar{z}_j(k-1) + \frac{1}{k} z_j(k) \quad (114-3)$$

$$\sigma_j^2 = \frac{k-1}{k} \sigma_j^2(k-1) + \frac{1}{k-1} (z_j(k) - \bar{z}_j(k))^2$$

شرایط اولیه بصورت $\bar{z}_j(1) = 0$ و $\sigma_j^2(1) = 0$ هستند.

فلوچارت الگوریتم شناسائی مدل eTS ساده‌سازی شده در صفحه بعد نشان داده شده است. در الگوریتم اصلی شناسائی (شکل ۳-۲) فاصله δ_{\min} روی بردار عناصر ورودی-خروجی (z) محاسبه می‌شد در نتیجه قواعد دارای مقدمه یکسان و نتیجه متفاوت می‌توانند در پایگاه وجود داشته باشند که تعبیر زبانی چنین قواعدی با تناقض مواجه می‌شود. اما در الگوریتم ساده‌سازی شده پارامتر δ_{\min} تنها روی بردار عناصر ورودی (x) محاسبه می‌شود تا از ایجاد مشکل مذکور جلوگیری شود.

نتایج اعمال الگوریتم مدل‌سازی ساده‌سازی شده به داده‌های سری مکی-گلاس و کوره گازی Box-Jenkins در جدول ۳-۱۱ دیده می‌شوند. از مقایسه این نتایج با جدول (۳-۴) نتیجه می‌شود که

اگر چه دقت الگوریتم ساده‌سازی شده از الگوریتم اصلی کمتر است اما روش ساده‌سازی شده نیز در مقایسه با بیشتر الگوریتم‌های مدل‌سازی برخط دیگر بهتر عمل می‌کند و در صورت استفاده از داده‌های زیاد، سریعتر بودن این الگوریتم در قیاس با روش اصلی نیز محسوس می‌باشد.

	سری مکی-گلاس	کوره گازی
NDEI	۰/۰۴۹۳	۰/۳۷۸
R	۵	۶

جدول ۱۱-۳- نتایج مدل‌سازی با استفاده از الگوریتم ساده‌سازی شده

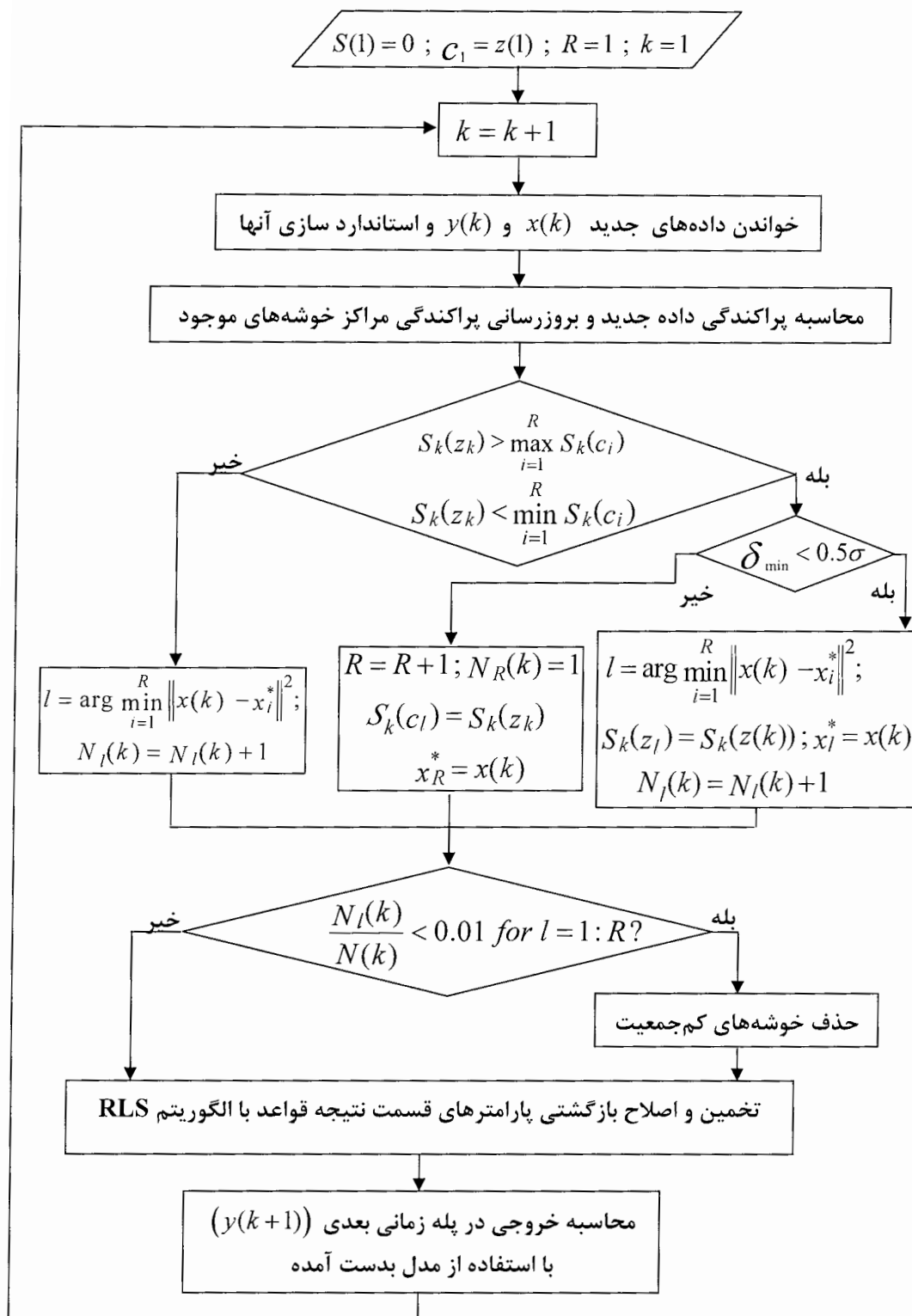
برای بررسی میزان مقاومت^۱ الگوریتم‌های مدل‌سازی ارائه شده در شکل‌های ۲-۳ و ۱۱-۳، ۵٪ نویز سفید گوسی به داده‌های سری مکی-گلاس اضافه کرده و الگوریتم‌های مذکور را به این داده‌ها اعمال می‌کنیم. نتایج حاصله در جدول ۱۲-۳ آورده شده‌اند.

	الگوریتم اصلی	الگوریتم ساده‌سازی شده
NDEI	۰/۵۴۸	۰/۶۸۳
R	۵	۷

جدول ۱۲-۳- بررسی مقاومت الگوریتم‌های مدل‌سازی در برابر نویز

ملاحظه می‌شود که مقاومت الگوریتم اصلی در برابر نویز بیشتر از الگوریتم ساده‌سازی شده است. همچنین تعداد قواعد در هر دو الگوریتم تنها باندازه یک قاعده نسبت به حالت بدون نویز افزایش یافته است که نشان‌دهنده مقاوم بودن هر دو الگوریتم خوشه‌سازی می‌باشد.

^۱ Robustness



شکل ۳-۱۲- الگوریتم شناسایی مدل eTS با اعمال ساده‌سازیهای محاسباتی

فصل چهارم

کنترل پیش‌بین‌تعمیم‌یافته فازی-تطبیقی

۴-۱- مقدمه

کنترل پیش‌بین فازی با معانی زیر در متن‌ها بکار رفته است :

۱. استفاده از مدل فازی بعنوان مدل پیش‌بین در MPC

۲. استفاده از توابع عضویت فازی برای بیان قیود متغیرها بصورت دقیق‌تر

۳. استفاده از قواعد فازی در بهینه‌سازی تابع هدف

بعلت اهمیت فوق‌العاده مدل پیش‌بین در عملکرد روشهای MPC و دقت بالاتر مدل‌های غیرخطی ، کنترلرهای MPC غیرخطی (مانند MPC فازی) در کنترل فرایندهای پیچیده موفق‌تر عمل می‌کنند با توجه به این مساله و ضعفهای موجود در مدل‌های پیش‌بین متداول ، در این پایان‌نامه از مدل فازی بعنوان مدل پیش‌بین در GPC استفاده می‌شود .

مرحله بحرانی موجود در کنترل MPC غیرخطی حداقل‌سازی تابع هزینه می‌باشد . در صورت استفاده از مدل LTI و عدم وجود قیود ، جواب تحلیلی برای مساله بهینه‌سازی وجود دارد همچنین بهنگام وجود قیود محدودکننده نیز مساله بهینه‌سازی غیرخطی و محدب بوده و به مساله برنامه‌ریزی درجه دوم^۱ (QP) تبدیل می‌شود که جواب تحلیلی ندارد اما با استفاده از روشهای عددی (مانند کاهش گرادیان سریع) می‌توانیم جواب بهینه کلی آنرا بصورت تضمینی بدست آوریم. در صورت استفاده از مدل غیرخطی TS (یا هر مدل غیرخطی دیگر) با یک مساله بهینه‌سازی غیرخطی (و غالباً غیرمحدب) مواجه می‌شویم که باید در هر پریود نمونه‌برداری حل شود. حل موثر این مساله با روشهای مبتنی بر تکرار در سیستمهای سریع با توجه به کوتاه بودن زمان نمونه‌برداری و حجم زیاد محاسبات امکانپذیر نمی‌باشد. گذشته از این الگوریتمهای بهینه‌سازی مبتنی بر تکرار نظیر روش ندر-مید^۲ یا روش برنامه‌ریزی درجه دوم ترتیبی (SQP^۳) غالباً به حداقل محلی همگرا می‌شوند که نتایج ضعیفی برای مساله بهینه‌سازی و

^۱ Quadratic Programming

^۲ Nedler-Mead

^۳ Sequential Quadratic Programming

کنترل بدست می‌دهد. با توجه به خاصیت خطی بودن محلی زیرمدلهای تشکیل‌دهنده مدل TS در این فصل دو روش برای حل مشکل بهینه‌سازی معرفی می‌شوند:

۱. استفاده از روش بهینه‌سازی Branch & Bound همراه کاهش ابعاد شبکه، گسترش الگوریتم استاندارد B&B از حالت MISO به حالت MIMO و تلاش برای کاهش حجم محاسبات.

۲. ترکیب کنترل GPC خطی با زیرمدلهای خطی محلی که با انجام "خطی‌سازی" از مدل غیرخطی فازی TS استخراج می‌شوند. البته واژه خطی‌سازی در اینجا نباید بصورت متداول خطی‌سازی محلی حول نقطه کار تعبیر شود همچنین اینکار بمعنی عدم آشکارسازی خواص غیرخطی سیستم نیست بلکه می‌توان مدل غیرخطی TS را در نواحی محلی بصورت یک مدل خطی با پارامترهای متغیر با زمان (وابسته به حالت‌های سیستم) در نظر گرفت. بسته به اینکه مدل مذکور در نقطه کاری حال حاضر یا در رشته‌ای از نقاط در پهنای افق پیش‌بینی بدست آمده باشد خطی‌سازی مذکور به دو نوع یک‌مرحله‌ای و چندمرحله‌ای تقسیم می‌شود که در ادامه توضیح داده خواهند شد.

۴-۲- فرمول‌بندی مدل فازی TS جهت تناسب با الگوریتمهای کنترل

یک سیستم MIMO با n_i متغیر ورودی و n_o متغیر خروجی در نظر می‌گیریم و این سیستم را با مجموعه‌ای از زیرمدلهای فازی TS گسسته MISO از نوع $NARX^1$ تقریب می‌زنیم. با معرفی دو عدد صحیح m و n (بطوریکه $n \leq m$) رشته مرتب‌شده از نمونه‌های تاخیری سیگنال y را بصورت زیر تعریف می‌کنیم:

$$\{y(k)\}_m^n \stackrel{def}{=} [y(k-m), y(k-m-1), \dots, y(k-m-n-1)] \quad (1-4)$$

خروجی زیرمدل l ام بصورت زیر می‌باشد:

$$y_l(k+1) = f_l(x_l(k)) \quad (2-4)$$

که بردار رگرسور $x_l(k)$ با رابطه زیر نمایش داده می‌شود:

¹ Nonlinear Auto-Regressive model with Exogenous inputs

$$x_l(k) = \left[\left\{ y_1(k) \right\}_0^{n_{y1}}, \dots, \left\{ y_{n_o}(k) \right\}_0^{n_{y/n_o}}, \left\{ u_1(k+1) \right\}_{nd1}^{n_{u1}}, \dots, \left\{ u_{n_i}(k+1) \right\}_{nd/n_i}^{n_{u/n_i}} \right] \quad (3-4)$$

در اینجا $[n_y]_{n_o \times n_o}$ و $[n_u]_{n_o \times n_i}$ ماتریسهائی هستند که تعداد نمونه‌های تاخیری ورودیها و خروجیها در آنها ذخیره شده است و $[n_d]_{n_o \times n_i}$ ماتریسی متشکل از مقادیر تاخیرهای خالص (انتقال) از ورودی به خروجی می‌باشد. قاعده i ام مدل فازی l ام را بصورت زیر بازنویسی می‌کنیم:

$$R_{li}: IF \ x_{l1}(k) \text{ IS } A_{li1} \text{ AND } \dots \text{ AND } x_{lp}(k) \text{ IS } A_{lip} \quad (4-4)$$

$$THEN \ y_{li}(k+1) = a_{li}y(k) + b_{li}u(k) + \theta_{li} \quad ; \quad i = 1, 2, \dots, K_l$$

در این رابطه x_{li} عنصری از بردار رگرسور، A_{li} مجموعه‌های فازی مقدمه قاعده i ام و θ بردار آفست می‌باشد. K_l نیز تعداد زیرمدلها (تعداد قواعد) مدل l ام را نشان می‌دهد. چندجمله‌ای‌های $a(z^{-1})$ و $b(z^{-1})$ بصورت زیر تعریف می‌شوند:

$$a_l(z^{-1}) = a_{0l} + a_{1l}z^{-1} + a_{2l}z^{-2} + \dots \quad (5-4)$$

$$b_l(z^{-1}) = b_{0l} + b_{1l}z^{-1} + b_{2l}z^{-2} + \dots$$

۳-۴- خطی‌سازی محلی مدل TS

۳-۴-۱- خطی‌سازی محلی تک‌مرحله‌ای مدل TS

جهت گسترش محدوده عملکرد کنترلرهای پیش‌بین خطی مدل TS فوق را بصورت محلی خطی-سازی کرده و بفرم فضای حالت LTV بازنویسی می‌کنیم. مجدداً یادآوری می‌شود که واژه خطی‌سازی در اینجا بمعنی خطی‌سازی ژاکوبین نیست بلکه منظور جایگذاری عناصر متغیر با زمان مدل فازی در مدل LTV است.

در کنترلر GPC خطی (بخش ۲-۶) از مدل خطی زیر جهت پیش‌بینی خروجی بصورت تابعی از سیگنال کنترل پیش‌بینی شده استفاده شده است:

$$x(k+1) = Ax(k) + Bu(k)$$

$$y(k) = Cx(k)$$

برای مدل فازی خطی‌سازی شده بصورت محلی معادلات فوق بفرم زیر درمی‌آیند :

$$x_{lin}(k+1) = x_{lin} + A^*(x_{lin}(k) - x_0) + B^*(u(k) - u_0) \quad (۶-۴)$$

$$y(k) = Cx_{lin}(k)$$

که x_0 و u_0 نقطه انجام خطی‌سازی را نشان می‌دهند. از مقایسه مدل فازی TS با روابط مشاهده می‌شود که بردار حالت‌های $x_{lin}(k)$ شامل بردارهای رگرسیون a_l متعلق به خروجیهای جداگانه است. در

هر لحظه ماتریسهای A^* ، B^* و C بدین‌صورت محاسبه می‌شوند : ابتدا آستانه‌های تحریک قواعد $\varphi_i(x_{lin}(k))$ محاسبه می‌شوند. سپس از مکانیزم استنتاج مدل فازی TS خواهیم داشت :

$$y_l(k+1) = \frac{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k)) y_{li}(k+1)}{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k))} \quad (۷-۴)$$

$$y_{li}(k+1) = a_{li} y(k) + b_{li} u(k) + \theta_{li} \quad (۸-۴)$$

اکنون a_l^* و b_l^* را بصورت زیر تعریف می‌کنیم :

$$a_l^* = \frac{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k)) a_{li}}{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k))} \quad (۹-۴)$$

$$b_l^* = \frac{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k)) b_{li}}{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k))} \quad (۱۰-۴)$$

$$\theta_l^* = \frac{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k)) \theta_{li}}{\sum_{i=1}^{K_l} \varphi_{li}(x_l(k))} \quad (۱۱-۴)$$

همچنین جهت نمایش فضای حالت بردارهای X ، u و y را بصورت زیر خواهیم داشت :

$$x(k) = \left[X_1(k), X_1(k-1), \dots, X_1(k-n_{y1}), \dots, X_{n_o}(k), X_{n_o}(k-1), \dots, X_{n_o}(k-n_{yn_o}) \right]^T \quad (12-4)$$

$$u(k) = \left[u_1(k-n_{d1}+1), u_1(k-n_{d1}), \dots, u_1(k-n_{d1}-n_{u1}+1), \dots, u_{n_i}(k-n_{d_{ni}}+1), u_{n_i}(k-n_{d_{ni}}), \dots, u_{n_i}(k-n_{d_{ni}}-n_{d_{uni}}+1) \right]^T \quad (13-4)$$

$$y(k) = \left[X_1(k), X_2(k), \dots, X_{n_o}(k) \right]^T \quad (14-4)$$

اکنون ماتریسهای مدل خطی محلی بفرم کانونی کنترل‌پذیر زیر بدست می‌آیند: A^* یک ماتریس

$$\cdot \alpha_1 = \sum_{j=1}^{n_o} n_{yj} \text{ است که } \alpha_1 \times \alpha_1$$

$$A^* = \begin{bmatrix} a_{11}^* & a_{12}^* & \dots & \dots & \dots & \dots & a_{1\alpha_1}^* & \theta_1^* \\ 1 & 0 & 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & 1 & \dots & \dots & \dots & \dots & \vdots & 0 \\ 0 & \dots & \ddots & \dots & \dots & \dots & \vdots & \vdots \\ a_{21}^* & a_{22}^* & \dots & \dots & \dots & \dots & a_{2\alpha_1}^* & \theta_2^* \\ 0 & \vdots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ a_{n_o 1}^* & a_{n_o 2}^* & \dots & \dots & \dots & \dots & a_{n_o \alpha_1}^* & \theta_{n_o}^* \\ 0 & \dots & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \dots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 & 0 & 1 \end{bmatrix} \quad (15-4)$$

$$\cdot \alpha_2 = \sum_{j=1}^{n_i} n_{uj} \text{ است که } \alpha_2 \times \alpha_2 \text{ با ابعاد } B^*$$

$$B^* = \begin{bmatrix} b_{11}^* & b_{12}^* & \dots & b_{1\alpha_2}^* \\ 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \\ b_{21}^* & b_{22}^* & & b_{2\alpha_2}^* \\ \vdots & \ddots & \ddots & \vdots \\ b_{n_o1}^* & b_{n_o2}^* & \dots & b_{n_o\alpha_2}^* \\ \vdots & \ddots & \ddots & \vdots \end{bmatrix} \quad (16-4)$$

C نیز دارای ابعاد $n_o \times \alpha_2$ می‌باشد :

$$C = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & & \ddots & & \ddots & & \vdots \\ 0 & \dots & \dots & 1 & 0 & \dots & 0 \end{bmatrix} \quad (17-4)$$

تابع هدف زیر را در نظر می‌گیریم :

$$J = \sum_{i=1}^{H_{2C}} \left\| w(k+i) - \hat{y}(k+i) \right\|_{P_i}^2 + \sum_{i=1}^{H_C} \left\| \Delta u(k+i-1) \right\|_{Q_i}^2 \quad (18-4)$$

با استفاده از ماتریسهای A^* ، B^* و C (بدون علامت ستاره) و تابع هدف فوق یک مساله MPC مقید خطی خواهیم داشت که با استفاده از برنامه درجه دوم زیر قابل حل می‌باشد :

$$\min_{\Delta \tilde{u}} \left\{ \frac{1}{2} \Delta \tilde{u}^T H \Delta \tilde{u} + c^T \Delta u \right\} \quad (19-4)$$

که در این رابطه

$$H = 2(R_u^T P R_u + Q) \quad (20-4)$$

$$c = 2 \left[R_u^T P^T (R_x A_k x(k) - w) \right]^T$$

قیود اعمال شده به u ، Δu و γ نیز با نامساوی زیر بیان می‌شوند :

$$\Lambda \Delta u \leq \omega \quad (21-4)$$

و داریم :

$$\Lambda = \begin{bmatrix} I_{\Delta u} \\ -I_{\Delta u} \\ I^{H_p m} \\ -I^{H_p m} \\ R_u \\ -R_u \end{bmatrix}, \quad \omega = \begin{bmatrix} u^{\max} - I_u u(k-1) \\ -u^{\min} - I_u u(k-1) \\ \Delta u^{\max} \\ -\Delta u^{\min} \\ y^{\max} - R_x A_k x(k) \\ -y^{\min} - R_x A_k x(k) \end{bmatrix} \quad (22-4)$$

که $I^{H_p m}$ یک ماتریس واحد با ابعاد $(H_p m \times H_p m)$ است. بقیه ماتریسها بصورت زیر تعریف می-

شوند :

$$R_x = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{H_p-1} \end{bmatrix}, \quad R_u = \begin{bmatrix} CB & 0 & \dots & 0 \\ CAB & CB & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{H_p-1}B & CA^{H_p-2}B & \dots & CA^{H_p-H_c}B \end{bmatrix} \quad (23-4)$$

رشته کنترلی بهینه $(\tilde{u}(k+i), i=1, \dots, H_c-1)$ بصورت زیر قابل محاسبه می‌باشد :

$$\begin{bmatrix} u(k) \\ \tilde{u}(k+1) \\ \vdots \\ \tilde{u}(k+H_c-1) \end{bmatrix} = \underbrace{\begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}}_{I_u} u(k-1) + \underbrace{\begin{bmatrix} I & 0 & \dots & 0 \\ I & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ I & I & \dots & I \end{bmatrix}}_{I_{\Delta u}} \begin{bmatrix} \Delta u(k) \\ \Delta \tilde{u}(k+1) \\ \vdots \\ \Delta \tilde{u}(k+H_c+1) \end{bmatrix} \quad (24-4)$$

در این حالت برخورد با محدودیت‌های اعمالی روی سیگنال کنترل آسان می‌باشد چون مساله بهینه‌سازی

QP بدست آمده با استفاده از تابع qp نرم‌افزار Matlab (جعبه‌ابزار بهینه‌سازی) قابل حل می‌باشد.

۴-۳-۲- خطی‌سازی محلی چندمرحله‌ای مدل TS

در خطی‌سازی تک‌مرحله‌ای در هر لحظه فرض می‌شود که مدل خطی با ماتریسهای بدست آمده در آن لحظه برای تمام افق پیش‌بینی معتبر می‌باشد. ممکن است مدل خطی مذکور نتواند رفتار سیستم غیرخطی در چندین پله زمانی جلوتر را با دقت مناسب پیش‌بینی کند که این امر بر عملکرد کنترلر تاثیر منفی می‌گذارد. برای حل این مشکل از مدل‌های خطی چندگانه $M(i)$ استفاده می‌شود که در امتداد منحنی عملکرد سیستم در افق پیش‌بینی بدست می‌آیند $(i = k, \dots, k + H_p)$. این فرایند خطی‌سازی چندمرحله‌ای نامیده می‌شود. در این روش استفاده از نمایش مدل فضای حالت LTV مشابه حالت خطی‌سازی تک‌مرحله‌ای (معادله) امکانپذیر نیست بلکه باید از مدل فضای حالت گسسته متغیر با زمان بصورت زیر استفاده شود:

$$\begin{aligned} x_{\text{lin}}(k+1) &= x_{\text{lin}}(k) + A_k \Delta x_{\text{lin}}(k) + B_k \Delta u(k) \\ y(k) &= C_k x_{\text{lin}}(k) \end{aligned} \quad (۲۵-۴)$$

$x_{\text{lin}}(k+1)$ بردار حالت پیش‌بینی شده بوده و A_k ، B_k و C_k نیز ماتریسهای متغیر با زمان مدل هستند. با نوشتن این فرمول‌بندی بصورت تکرارشونده می‌توانیم خروجیهای آینده فرایند را پیش‌بینی کنیم. در هر لحظه زمانی k بردارهای حالت و تغییرات سیگنال کنترل معلوم هستند بنابراین حالت‌های آینده بصورت زیر بدست می‌آیند:

۱. استفاده از مدل خطی $M(k)$ جهت محاسبه سیگنال کنترل $u(k)$ روی تمام افق پیش‌بینی
۲. محاسبه $\hat{y}(k+1)$ با استفاده از $u(k)$
۳. خطی‌سازی محلی مدل TS حول $(\hat{y}(k+1), u(k))$ و بدست آوردن $M(k+1)$
۴. محاسبه رشته کنترلی جدید $u(k+1)$ روی تمام افق پیش‌بینی با استفاده از $M(k)$ و $M(k+1)$
۵. استفاده از $u(k)$ و $u(k+1)$ جهت محاسبه $\hat{y}(k+2)$
۶. خطی‌سازی محلی مدل TS حول $(\hat{y}(k+2), u(k+1))$ و بدست آوردن $M(k+2)$

۷. محاسبه رشته کنترلی جدید $u(k+2)$ روی تمام افق پیش‌بینی با استفاده از $M(k)$ ، $M(k+1)$ و

$$M(k+2)$$

۸. تکرار مراحل (۵-۷) برای $i = k+2 : k+H_{2P}$ و محاسبه سیگنال کنترول نهائی بر اساس

$M(k)$ ، $M(k+1)$ ، ...، $M(k+H_{2P})$. مساله بهینه‌سازی QP برای حالت خطی‌سازی چندمرحله-

ای بصورت زیر خواهد بود :

$$\min_{\Delta \tilde{u}} \left\{ \frac{1}{2} \Delta \tilde{u}^T H \Delta \tilde{u} + c^T \Delta u \right\} \quad (۲۶-۴)$$

که داریم :

$$H = 2(R_u^T P R_u + Q) \quad (۲۷-۴)$$

$$c = 2[R_u^T P^T (R_x A_k x(k) - w)]^T$$

همچنین قيود اعمالی روی u ، Δu و y بصورت زیر بیان می‌شوند :

$$\Lambda \Delta u \leq \omega \quad (۲۸-۴)$$

$$\Lambda = \begin{bmatrix} I_{\Delta u} \\ -I_{\Delta u} \\ I^{H_{2P}m} \\ -I^{H_{2P}m} \\ R_u \\ -R_u \\ R_{u1} \\ dR_{u1} \\ -R_{u1} \\ -dR_{u1} \end{bmatrix}, \quad \omega = \begin{bmatrix} u^{\max} - I_u u(k-1) \\ -u^{\min} - I_u u(k-1) \\ \Delta u^{\max} \\ -\Delta u^{\min} \\ y^{\max} - R_x A_k x(k) \\ -y^{\min} - R_x A_k x(k) \\ dy^{\max} - R_{x1} A_k x(k) + y(k) \\ dy^{\max1} - d R_x A_k x(k) \\ -dy^{\min} + R_{x1} A_k x(k) \\ -dy^{\min1} + d R_x A_k x(k) \end{bmatrix} \quad (۲۹-۴)$$

در این رابطه نیز ابعاد ماتریس واحد $I^{H_p m}$ بصورت $(H_p m \times H_p m)$ است. برای رشته کنترل بهینه نیز داریم:

$$\begin{bmatrix} u(k) \\ \tilde{u}(k+1) \\ \vdots \\ \tilde{u}(k+H_c-1) \end{bmatrix} = \underbrace{\begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}}_{I_u} u(k-1) + \underbrace{\begin{bmatrix} I & 0 & \dots & 0 \\ I & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ I & I & \dots & I \end{bmatrix}}_{I_{\Delta u}} \begin{bmatrix} \Delta u(k) \\ \Delta \tilde{u}(k+1) \\ \vdots \\ \Delta \tilde{u}(k+H_c+1) \end{bmatrix} \quad (30-4)$$

$$R_x = \begin{bmatrix} C_k \\ C_{k+1}A_k \\ \vdots \\ C_{k+H_p}A_{k+H_p} \dots A_k \end{bmatrix} \quad (31-4)$$

$$R_u = \begin{bmatrix} C_{k+1}B_k & \dots & 0 \\ C_{k+2}A_{k+1}B_{k+1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ C_{k+H_p}A_{k+H_p} \dots A_{k+1}B_k & \dots & C_{k+H_p}A_{k+H_p-H_c} \dots A_{k+1}B_{k+H_c} \end{bmatrix} \quad (32-4)$$

$$dR_x = \begin{bmatrix} C_{k+1}A_k - C_k \\ C_{k+2}A_{k+1}A_k - C_{k+1}A_k \\ M \\ C_{k+H_p} \prod_{i=H_p-1}^0 A_{k+i} - C_{k+H_p-1} \prod_{i=H_p-2}^0 A_{k+i} \end{bmatrix} \quad (33-4)$$

$$dR_u = \begin{bmatrix} C_{k+2} A_{k+1} B_k - C_{k+1} B_k & \dots \\ C_{k+3} A_{k+2} A_{k+1} B_k - C_{k+2} A_{k+1} B_k & \dots \\ \vdots & \ddots \\ C_{k+H_p} \left(\prod_{i=H_p-1}^1 A_{k+i} \right) B_k - C_{k+H_p-1} \left(\prod_{i=H_p-2}^1 A_{k+i} \right) B_k & \dots \\ 0 \\ 0 \\ \vdots \\ C_{k+H_p} \left(\prod_{i=H_p-1}^{H_C+1} A_{k+i} \right) B_k - C_{k+H_p-1} \left(\prod_{i=H_p-2}^{H_C+1} A_{k+i} \right) B_{k+H_C} \end{bmatrix} \quad (34-4)$$

۴-۳-۳- رابطه خطی سازی محلی با خطی سازی ژاکوبین

رابطه ورودی- خروجی مدل TS را مجدداً یادآوری می‌کنیم:

$$y_1(k+1) = \frac{\sum_{i=1}^{K_l} \varphi_{li}(y(k), u(k)) (a_{li} y(k) + b_{li} u(k) + \theta_{li})}{\sum_{i=1}^{K_l} \varphi_{li}(y(k), u(k))} \quad (35-4)$$

$$= \sum_{i=1}^{K_l} \lambda_{li}(y(k), u(k)) \cdot f_{li}(y(k), u(k))$$

در این رابطه λ_{li} آستانه تحریک نرمالیزه شده قواعد بوده و بصورت زیر بدست می‌آید:

$$\lambda_{li}(y(k), u(k)) = \frac{\varphi_{li}(y(k), u(k))}{\sum_{i=1}^{K_l} \varphi_{li}(y(k), u(k))} \quad (36-4)$$

همچنین f_{li} زیرمدل خطی i امین قاعده مدل l می‌باشد:

$$f_{li}(y(k), u(k)) = a_{li} y(k) + b_{li} u(k) + \theta_{li} \quad (37-4)$$

با اعمال خطی سازی ژاکوبین به رابطه (۳۵-۴) داریم:

$$\frac{\partial y_l(k+1)}{\partial y(k)} = \sum_{i=1}^{K_l} \left(\frac{\partial \varphi_{li}(y(k), u(k))}{\partial y(k)} \cdot f_{li}(y(k)) + \varphi_{li}(y(k), u(k)) \frac{\partial f_{li}(y(k))}{\partial y(k)} \right) \quad (38-4)$$

$$\frac{\partial y_l(k+1)}{\partial u(k)} = \sum_{i=1}^{K_l} \left(\frac{\partial \varphi_{li}(y(k), u(k))}{\partial u(k)} \cdot f_{li}(u(k)) + \varphi_{li}(y(k), u(k)) \frac{\partial f_{li}(u(k))}{\partial u(k)} \right) \quad (39-4)$$

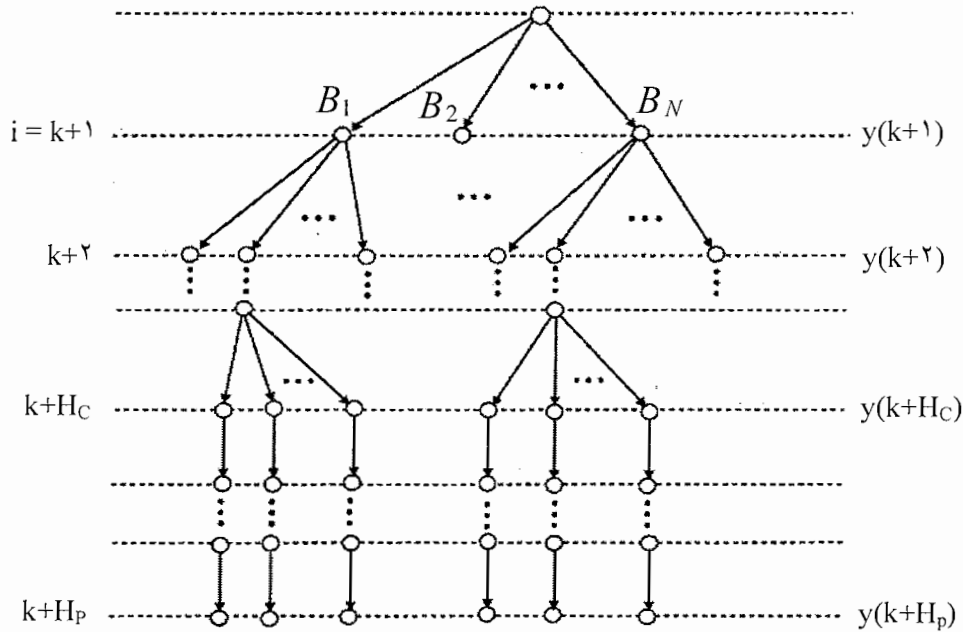
از مقایسه این نتایج با روابط ۸-۴ تا ۱۱-۴ دیده می‌شود که جملات شامل مشتق آستانه‌های تحریک نرمالیزه در عبارات ۸-۴ تا ۱۱-۴ وجود ندارند. گذشته از تفاوت‌های موجود با روش خطی‌سازی استاندارد، روش خطی‌سازی مدل TS یک مدل خطی با پارامترهای متغیر با زمان بدست می‌دهد که بخوبی قادر به ردیابی رفتار سیستم غیرخطی می‌باشد. یکی از مزایای عمده روش خطی‌سازی ارائه شده اینست که احتیاجی به محاسبات (احتمالاً وقت‌گیر) مشتق‌گیری ندارد. در قسمت شبیه‌سازی فصل پنجم برتری روش ارائه شده نسبت به خطی‌سازی ژاکوبین بیشتر مشخص خواهد شد.

۴-۴- روش بهینه‌سازی B&B

مساله بهینه‌سازی MPC غیرخطی را می‌توان بصورت مساله جستجو در فضای گسسته سیگنال کنترل فرمول‌بندی کرد. روش B & B با استفاده از ساختار درختی^۱ مساله بهینه‌سازی را به چندین زیرمساله تقسیم کرده و آنرا حل می‌کند و متکی بر این نکته است که در فضای ناحیه‌بندی شده جوابها در هر لحظه تنها تعداد کمی از نواحی (جوابها یا شاخه‌ها) باید آزمایش شوند و از بقیه نواحی موجود در فضای سیگنال کنترل صرف‌نظر می‌شود چون حاوی جواب بهینه نمی‌باشند. شکل ۴-۱ اصول روش B & B را نشان می‌دهد.

نواحی که حاوی جواب بهینه هستند مجدداً به نواحی (زیرشاخه‌های) کوچکتری تقسیم می‌شوند. با تعریف باندهای بالا و پائین برای تابع هدف تصمیم می‌گیریم که عمل تقسیم شاخه و امتحان زیرشاخه‌ها برای کدام شاخه‌ها ادامه پیدا کند.

^۱ Tree Structure



شکل ۴-۱- دیاگرام شماتیک روش بهینه‌سازی B & B

ورودی کنترل مقادیری از مجموعه گسسته B را اختیار می‌کند:

$$B = \{u_1 \times \dots \times u_{ni}\} \quad (40-4)$$

$$U_i = \{u_{i,\min}, \dots, u_{i,\max}\} \quad (41-4)$$

تعداد عناصر U_i عدد گسسته‌سازی (N_i) نامیده می‌شود. معمولاً N_i مقدار عددی بطور یکنواخت در فضای U_i پخش شده‌اند. i بیانگر ارتفاع درخت بوده و r شاخه‌ای از درخت می‌باشد که بیانگر انتخاب B_j می‌باشد. در ارتفاع i بتعداد N^{ni} انتخاب برای سیگنال کنترل داریم که حداکثر باعث بدست آمدن N^{ni} شاخه می‌شود. اگر همه شاخه‌های موجود در هر ارتفاع به زیرشاخه‌هایی تقسیم شوند آنگاه باید حالت را مورد بررسی قرار دهیم که بار محاسباتی خیلی زیادی ایجاد خواهد کرد. برای حل $(N^{ni})^{Hc}$

این مشکل باندهای بالا و پائین را به تابع هدف اعمال می‌کنیم. بدین منظور تابع هدف تجمعی^۱ $J(i)$ بصورت زیر تعریف می‌شود:

$$J(i) = \sum_{m=1}^i \|w(k+m) - \hat{y}(k+m)\|_{p_m}^2 + \sum_{m=1}^i \|\Delta u(k+m-1)\|_{Q_m}^2 \quad (42-4)$$

تقسیم شاخه r ام موجود در ارتفاع i ام به زیرشاخه‌ها و آزمایش آنها در تابع هدف در صورتی انجام می‌شود که تابع هزینه تجمعی باضافه هزینه موجود از ارتفاع i تا انتهای افق پیش‌بینی از باند بالایی هزینه کلی (J_U) کمتر باشد. هزینه موجود بین ارتفاع i تا انتهای افق پیش‌بینی عموماً نامعلوم می‌باشد اما می‌توان آنرا بصورت مجموع دو جمله نوشت: یکی $J_r(i)$ که متناظر با انتقال (r, B) بوده و با ارزیابی عنصر متناظر در تابع هدف $J(i)$ محاسبه می‌شود. جمله دوم نیز باند پائین تخمینی برای هزینه در پله‌های زمانی باقیمانده $H_p, \dots, i+1$ است که با $J_L(i+1)$ نمایش داده می‌شود. بنابراین شرط تقسیم شاخه r ام در ارتفاع i ام باین شکل بیان خواهد شد:

$$J(i) + J_r(i) + J_L(i+1) < J_U \quad (43-4)$$

محاسبه $J_L(i+1)$ بسیار مشکل می‌باشد و بدون ایجاد مشکل خاصی می‌توانیم آنرا صفر فرض کنیم. توجه شود که برای $i > H_C - 1$ هیچگونه شاخه سازی جدیدی انجام نمی‌شود و آخرین عنصر رشته کنترل $(u(k + H_C - 1))$ از H_C تا H_P به مدل اعمال می‌شود.

در مورد باندهای اعمالی به تابع هدف می‌توان گفت که باند بالا باید تا حد امکان کوچک بوده و باند پائین تا حد امکان بزرگ باشد تا تعداد شاخه‌ها و در نتیجه فضای جستجو کاهش یابد. اگر مکانیزمی برای تعیین باندها وجود نداشته باشد باند بالا را در ابتدا بی‌نهایت می‌گیریم. اولین مسیر در جستجوی درختی با استفاده از استراتژی شبکه‌بندی انتخاب کمترین مقدار $J_r(i)$ در هر ارتفاع i تعیین می‌شود. در صورتی که منحنی مرجع ثابت بوده یا دارای تغییرات آهسته باشد باند بالایی برابر تابع هزینه ترمینال در نظر گرفته می‌شود $(J_U = J(H_P))$. اگر در ارتفاع بالاتری از درخت جستجو، $J(H_{2P}) < J_U$ ،

^۱ Cumulative

یافته شود آنگاه J_u جدید را برابر $J(H_{2P})'$ در نظر می‌گیریم. در صورت نیافتن تخمین بهتر باند پائینی $J_L(i)$ برابر صفر فرض می‌شود ($i=1,2,\dots,H_C-1$). تجربیات بدست آمده در مورد این الگوریتم نشان می‌دهند که حتی بدترین تخمین‌های در نظر گرفته شده برای باندهای بالا و پائین بصورت فوق نیز می‌توانند تا حدود زیادی باعث کاهش حجم محاسبات و فضای جستجو شوند.

روش بهینه‌سازی $B \& B$ بکار رفته در کنترل پیش‌بین نسبت به روشهای بهینه‌سازی غیرخطی دارای سه مزیت عمده می‌باشد:

۱. جواب بدست آمده غالباً بهینه عمومی است در نتیجه عملکرد بهینه کنترلر بدون در نظر گرفتن هیچگونه شرطی روی شکل تابع هدف تضمین می‌شود.
۲. حدس اولیه‌ای برای شروع الگوریتم مورد نیاز نیست بنابراین برخلاف روشهای بهینه‌سازی مبتنی بر تکرار نحوه شروع الگوریتم نمی‌تواند روی نتیجه نهائی تاثیر منفی داشته باشد.
۳. روش $B \& B$ بطور صریح از قیود اعمالی به متغیرها در تعیین سیگنال کنترل استفاده می‌کند و در حقیقت وجود این قیود باعث محدود شدن فضای جستجو و در نتیجه افزایش بازدهی فرایند تقسیم شاخه‌ها می‌شود. همچنین در تعداد زیادی از سیستمهای صنعتی متغیرهای کنترلی تنها می‌توانند مقادیر گسسته داشته باشند که روش $B \& B$ می‌تواند باسانی در تعیین اینگونه سیگنالها بکار برده شود در حالیکه روشهای بهینه‌سازی عددی گرادینانی در مورد این سیستمها با مشکل مواجه می‌شوند.

روش $B \& B$ دارای دو عیب ذاتی است: یکی اینکه حجم محاسبات با تعداد تقسیمات فضای جستجو و افق کنترل رابطه نمائی دارد. در مورد سیستمهای $MIMO$ فضای جستجو بسرعت بزرگ می‌شود بنابراین روش $B \& B$ برای سیستمهای با تعداد متغیرهای ورودی-خروجی زیاد یا افق کنترل بزرگ قابل استفاده نیست. عیب دوم اینست که سیگنال کنترل تنها مقادیر گسسته محدودی می‌تواند داشته باشد. با افزایش تعداد این مقادیر کیفیت جواب نهائی بهتر شده و از طرف دیگر حجم محاسبات نیز زیاد می‌شود در نتیجه باید مصالحه‌ای بین تعداد تقسیمات و کیفیت عملکرد سیستم کنترل برقرار شود.

برای کاهش اثر خواص نامطلوب ذکر شده چندین راه‌حل در مراجع مختلف پیشنهاد شده‌اند که در اینجا برخی از این راه‌حلها باختصار توضیح داده می‌شوند:

۱. اصلاح بازگشتی تعداد تقسیمات شاخه‌ها، باین صورت که در ابتدای درخت جستجو، تعداد شاخه‌ها کم انتخاب می‌شود تا بار محاسباتی کمتر شود. با افزایش ارتفاع درخت و نزدیک شدن به جواب بهینه، تعداد تقسیمات (شاخه‌ها) نیز افزایش می‌یابد. افزایش تعداد تقسیمات باعث هموارتر شدن سیگنال کنترل شده و از نوسان سیستم بین مقادیر گسسته سیگنال کنترل جلوگیری می‌کند.

۲. جستجوی شبکه‌ای غالباً باعث کاهش محاسبات می‌شود. بعد از هر شاخه، تابع هزینه $J(i)$ برای تمام انتخابهای کنترلی ممکن در این مرحله محاسبه می‌شود و هزینه‌های جزئی مذکور بر حسب اندازه مرتب می‌شوند بطوریکه انتخابهای کنترلی دارای کمترین هزینه بیشترین شانس را برای شاخه‌سازی جدید خواهند داشت.

۳. در صورتی که تغییرات سیگنال مرجع آهسته باشد آخرین رشته سیگنال کنترلی $u(k, \dots, k + H_p - 1)$ محاسبه شده در زمان $k-1$ تخمین خوبی برای رشته کنترلی بهینه در زمان k است. خروجی $y(k+1)$ با استفاده از ورودی تخمینی بسرعت قابل محاسبه بوده و هزینه بدست آمده بعنوان باند بالای هزینه در نظر گرفته می‌شود.

۴. اندازه دینامیک شبکه (سوسا و دیگران، [۴۴]) از یک مجموعه تطبیقی سیگنالهای کنترل گسسته بر اساس یک معیار فازی ساده استفاده می‌کند. عمل تطبیق با ضرب مجموعه دینامیک اعمال کنترلی در ضریب مقیاس‌بندی^۱ حاصل می‌شود. معیار فازی از سیگنال خطا و تغییرات آن برای یافتن ضریب مقیاس‌بندی استفاده می‌کند. راه حل فوق در [۴۵] به حالت MIMO گسترش داده شده است.

^۱ Scaling Factor

۴-۵- الگوریتم GPC فازی-تطبیقی با استفاده از خطی‌سازی مدل TS

اگر در الگوریتم GPC فازی-تطبیقی از خطی‌سازی محلی مدل TS استفاده شود بطور همزمان می‌توانیم از مزایای مدل‌سازی فازی و مزایای وجود جواب تحلیلی برای مساله بهینه‌سازی در حالت GPC خطی استفاده کنیم و برتری عمده کاربرد مدل فازی TS نسبت به دیگر مدل‌های فازی نظیر مدل‌های رابطه‌ای^۱ در همین نکته نهفته است. الگوریتم GPC فازی-تطبیقی از خطی‌سازی محلی مدل TS برای کنترل یک فرایند ناشناخته شامل انجام مراحل زیر در هر لحظه می‌باشد :

۱. اندازه‌گیری متغیرهای ورودی - خروجی سیستم و فازی‌سازی آنها جهت محاسبه سطوح فعال‌سازی نرمالیزه شده قواعد فازی

۲. بروزرسانی برخط پارامترهای زیرمدل‌های خطی مدل فازی-تطبیقی TS با استفاده از داده‌های جدید. در این مرحله با توجه به توان محاسباتی در دسترس، دقت مورد نیاز و سرعت سیستم تحت کنترل، یکی از الگوریتم‌های شناسایی سیستم برخط مدل فازی-تطبیقی TS (شکل ۳-۲ یا ۳-۱۱) را بکار می‌بریم.

۳. خطی‌سازی مدل فازی-تطبیقی TS بصورت محلی با استفاده از سطوح فعال‌سازی نرمالیزه شده قواعد (λ_i) و بردار ورودی گسترش یافته (محاسبه بردار پارامترهای وزن‌دار مدل، ψ)

۴. بدست آوردن ماتریس‌های مدل فضای حالت خطی $\{A, B, C\}_k$ که بطور لحظه‌ای بر سیستم واقعی منطبق می‌باشند.

۵. ترکیب مدل فضای حالت خطی $\{A, B, C\}_k$ با روابط GPC خطی (فضای حالت) و محاسبه رشته قانون کنترل بهینه.

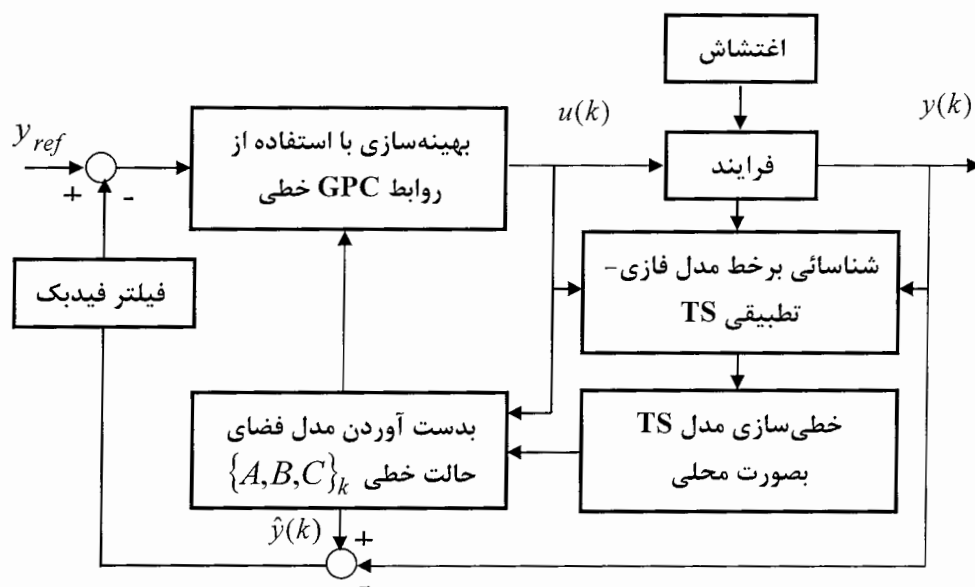
۶. اعمال اولین عنصر از رشته کنترلی بهینه به سیستم و بازگشت به مرحله اول.

MPC در اصل یک استراتژی کنترل حلقه باز است اما با استفاده از مفهوم افق پس‌رونده نوعی فیدبک حلقه‌بسته ایجاد می‌شود. بهر صورت جهت کاهش تاثیر اغتشاش، نویز اندازه‌گیری و خطای پیش-

^۱ Relational

بینی ناشی از عدم تطابق مدل-سیستم^۱ عموماً از طرح IMC^2 برای پیاده‌سازی الگوریتم MPC استفاده می‌کنند.

این طرح شامل سه قسمت می‌باشد: مدلی که اثر سیگنال کنترل روی سیستم را پیش‌بینی می‌کند، کنترلر بهینه و فیلتر فیدبک. این فیلتر باعث افزایش مقاومت در برابر اغتشاش، نویزهای اندازه‌گیری و عدم مطابقت مدل-سیستم شده و با کاهش دادن بهره حلقه، پایداری سیستم را زیاد می‌کند. دیاگرام بلوکی پیاده‌سازی الگوریتم GPC فازی-تطبیقی با استفاده از خطی‌سازی محلی مدل TS به همراه طرح IMC در شکل زیر نشان داده شده است.



شکل ۴-۲- پیاده‌سازی الگوریتم GPC فازی-تطبیقی با استفاده از خطی‌سازی محلی مدل TS به همراه طرح IMC

^۱ Model/System Mismatch

^۲ Internal Model Control

فصل پنجم

کنترل CSTR

با روش

GPC فازي - تطبیقی

۵-۱- مقدمه

در این فصل ابتدا با ارائه چند مثال و استفاده از قوانین اساسی مدلسازی سیستمهای شیمیائی، مدل ریاضی سیستم CSTR را برای چند حالت مختلف بدست می‌آوریم و مشخصات CSTR بکار رفته در پایان‌نامه را بیان می‌کنیم سپس روش GPC فازی به سیستم اعمال می‌شود و نتایج بدست آمده با نتایج کنترل سیستم با روش GPC خطی مقایسه می‌شوند. همچنین مقاومت GPC فازی در برابر تغییر پارامترهای سیستم و وجود نویز اندازه‌گیری مورد بحث و بررسی قرار می‌گیرد.

۵-۲- معرفی سیستم CSTR

۵-۲-۱- قوانین اساسی

تعدادی از قوانین اساسی که در مدلسازی سیستمها کاربرد دارند عبارتند از معادلات پیوستگی (معادله پیوستگی کلی (تعادل جرم) و معادله پیوستگی اجزاء (تعادل اجزاء))، معادله انرژی، معادلات حرکت، انتقال و حالت، معادلات تعادل (تعادل شیمیائی و تعادل فاز) و سینتیک شیمیائی. نحوه کاربرد بعضی از قوانین فوق در مدلسازی ریاضی CSTR با چند مثال توضیح داده می‌شود.

مثال ۵-۱ (سیستم CSTR بدون خنک‌کننده): یک تانک مایع اختلاط کامل^۱ که در آن واکنش شیمیائی رخ می‌دهد را در نظر بگیرید. دبی مایع ورودی با F_0 و چگالی آن با ρ_0 نشان داده می‌شود. همچنین حجم راکتور، غلظت ماده A در مایع ورودی به سیستم و غلظت آن در درون راکتور به ترتیب با V ، C_{A0} و C_A نمایش داده شده‌اند. فرض کنیم یک واکنش مرتبه اول ساده در این تانک رخ می‌دهد. نرخ مصرف واکنش‌دهنده A در واحد حجم^۲ با غلظت لحظه‌ای A در درون تانک ارتباط مستقیمی خواهد داشت. با علامت‌گذاریهای فوق، جریان ماده A به درون سیستم برابر $F_0 C_{A0}$ و جریان آن بسمت بیرون مساوی $F C_A$ خواهد شد. همچنین نرخ تشکیل ماده A در اثر واکنش بصورت $kV C_A$ - درمی-

^۱. Perfectly Mixed Tank

^۲. Per Unit Volume

آید که پارامتر k بیانگر سرعت واکنش بوده و واحد آن عکس زمان است و علامت منفی هم مصرف شدن ماده A را نشان می‌دهد. همچنین داریم:

$$\text{نرخ تغییر ماده } A \text{ در درون تانک} = \frac{d}{dt}(V C_A) \quad (1-5)$$

با ترکیب روابط فوق، معادله پیوستگی اجزاء برای واکنش‌دهنده A بدین شکل خواهد بود:

$$\frac{d}{dt}(V C_A) = F_0 C_{A0} - F C_A - kV C_A \quad (2-5)$$

همچنین از معادله پیوستگی کلی (تبادل جرم) داریم:

$$\frac{d(PV)}{dt} = F_0 \rho_0 - F \rho \quad (3-5)$$

مثال ۲-۵) سیستم CSTR با خنک‌کننده: برای خارج کردن گرمای واکنش (λ)، به سیستم قبلی یک کوپل خنک‌کننده اضافه می‌کنیم. فرض شده است λ برای واکنش گرماده منفی و برای واکنش گرماگیر مثبت باشد. نرخ تولید حرارت در اثر واکنش باینصورت خواهد بود:

$$Q_G = -\lambda V C_A k \quad (4-5)$$

درجه حرارت مواد ورودی و مواد درون راکتور بترتیب T_0 و T است. معادله انرژی برای سیستم مذکور باینصورت خواهد بود:

$$F_0 \rho_0 (U_0 + K_0 + \phi_0) - F \rho (U + K + \phi) + (Q_G + Q) - (W + F P + F_0 P_0) \frac{1}{J} = \frac{d}{dt} [(U + K + \phi) V \rho]; \quad (5-5)$$

در عبارت فوق U انرژی داخلی، K انرژی جنبشی، ϕ انرژی پتانسیل، W کار انجام شده، P فشار سیستم، P_0 فشار جریان ورودی به سیستم و J یک ثابت است. اگر سرعت جریانهای محوری خیلی زیاد نباشد انرژی جنبشی قابل اغماض خواهد بود. همچنین در صورتیکه محللهای ورود و خروج مواد

اختلاف ارتفاع زیادی نداشته باشند از انرژی پتانسیل نیز صرفنظر می‌شود بنابراین معادله (۵-۱) بفرم کاهش یافته زیر درمی‌آید :

$$\begin{aligned} \frac{d(\rho V U)}{dt} &= F_0 \rho_0 U_0 - F \rho U + Q_G + Q - F \rho \left(\frac{P}{\rho} \right) \frac{1}{J} + F_0 \rho_0 \left(\frac{P_0}{\rho_0} \right) \frac{1}{J} \\ &= F_0 \rho_0 \left(U_0 + \frac{P_0 \bar{V}_0}{J} \right) - F \rho \left(U + \frac{P \bar{V}}{J} \right) + Q_G + Q \end{aligned} \quad (6-5)$$

\bar{V} حجم مخصوص (معکوس چگالی) است . آنتالپی بصورت زیر تعریف می‌شود:

$$h \text{ or } H = U + P \bar{V} / J \quad (7-5)$$

ما از علامتهای h و H بترتیب برای آنتالپی جریانهای مایع و بخار استفاده می‌کنیم . بنابراین معادله (۵-۶) برای سیستم CSTR بصورت زیر درمی‌آید :

$$\frac{d(\rho V U)}{dt} = F_0 \rho_0 h_0 - F \rho h + Q - \lambda V k C_A \quad (8-5)$$

در تعداد زیادی از سیستمها جمله $P \bar{V}$ در مقایسه با V قابل صرفنظر کردن است و از نرخ زمانی تغییر آنتالپی سیستم بجای انرژی داخلی سیستم استفاده می‌شود :

$$\frac{d(\rho V h)}{dt} = F_0 \rho_0 h_0 - F \rho h + Q - \lambda V k C_A \quad (9-5)$$

از ترمودینامیک می‌دانیم که ظرفیت حرارتی در فشار ثابت (c_p) و در حجم ثابت (c_v) بصورت زیر تعریف می‌شوند :

$$c_p = \left(\frac{\partial H}{\partial T} \right)_P ; \quad c_v = \left(\frac{\partial H}{\partial T} \right)_V \quad (10-5)$$

برای اینکه نشان دهیم معادله انرژی بوسیله درجه حرارت تحت تاثیر قرار می‌گیرد اجازه دهید برای سادگی کار آنتالپی مایع را بصورت حاصلضرب درجه حرارت مطلق در ظرفیت حرارتی میانگین c_p (که عدد ثابتی است) بیان کنیم :

$$h = c_p T \quad (11-5)$$

یک فرض ساده‌کننده دیگر نیز بکار می‌بریم: چگالی مایع را ثابت و انرژی داخلی را تنها مساوی $c_p T$ در نظر می‌گیریم. با این ساده‌سازیها معادله (۹-۵) بصورت زیر درمی‌آید:

$$\rho c_p \frac{d(VT)}{dt} = \rho c_p (F_0 T_0 - F \rho) + Q - \lambda V k C_A \quad (12-5)$$

۵-۲-۲- سینیتیک شیمیائی

برای مدلسازی راکتورها باید با روابط اولیه توصیف‌کننده سینیتیک واکنشهای شیمیائی آشنا باشیم بعنوان نمونه‌ای از این روابط، وابستگی حرارتی آرنیوس^۱ را بیان می‌کنیم. این رابطه می‌گوید که وابستگی حرارتی یک سرعت واکنش خاص معمولاً بصورت نمائی است:

$$k = \alpha \cdot \exp\left(-\frac{E}{RT}\right) \quad (13-5)$$

که E انرژی فعالسازی، T درجه حرارت، R ثابت گاز کامل و α یک عدد ثابت است. وابستگی نمائی موجود در رابطه سرعت واکنش یکی از مهمترین و شدیدترین غیرخطی‌های موجود در سیستمهای شیمیائی است.

۵-۲-۳- CSTR غیرهمدم

در این حالت یک واکنش گرماده برگشت‌ناپذیر در راکتور اختلاط کامل رخ داده و ماده A به ماده B تبدیل می‌شود. واکنش نسبت به ماده A از مرتبه اول بوده و λ گرمای واکنش را نشان می‌دهد. از تلفات حرارتی صرف‌نظر شده و چگالی‌ها ثابت فرض می‌شوند.

برای انتقال حرارت واکنش به بیرون راکتور یک ژاکت خنک‌کننده راکتور را احاطه کرده است. آب خنک‌کننده با نرخ F_r و درجه حرارت اولیه T_{r0} وارد ژاکت می‌شود. حجم آب درون ژاکت (V_r) ثابت در نظر گرفته می‌شود. معادله تبادل حرارت بین فرایند و آب خنک‌کننده بصورت زیر است:

$$Q = V_r \cdot A_H (T - T_r) \quad (14-5)$$

^۱. Arrhenius Temperature Dependence

در این رابطه Q نرخ انتقال حرارت، V ثابت انتقال حرارت و A_H ناحیه انتقال حرارت است. سایر معادلات توصیف کننده دینامیک سیستم از این قرار می باشند:

$$\frac{dV}{dt} = F_0 - F \quad \text{معادله پیوستگی کلی راکتور}$$

$$\frac{d}{dt}(V C_A) = F_0 C_{A0} - F C_A - V k C_A^n \quad \text{معادله پیوستگی اجزاء راکتور}$$

$$\rho \frac{d}{dt}(Vh) = \rho(F_0 h_0 - F h) - \lambda V k C_A^n - V A_H (T - T_J) \quad \text{معادله انرژی راکتور}$$

$$\rho_J V_J \frac{d h_J}{dt} = F_J \rho_J (h_{J0} - h_J) + V A_H (T - T_J) \quad \text{معادله انرژی ژاکت}$$

پارامترهای ρ ، h ، ρ_J و h_J به ترتیب چگالی و آنتالپی مایع فرایند و آب خنک کننده را نشان می دهند. با فرض ثابت بودن چگالی، C_p و C_v مساوی می شوند و می توانیم از آنتالپی در مشتقات زمانی استفاده کنیم تا جایگزین انرژیهای داخلی شود.

رابطه هیدرولیکی مابین حجم مایع درون راکتور و جریان خروجی از آن نیز مورد نیاز است. فرض می کنیم که یک کنترلر تناسبی سطح مایع $(F = K_V (V - V_{\min}))$ ، نرخ جریان خروجی را متناسب با حجم مایع درون تانک تنظیم کند. V_{\min} حداقل مقدار مجاز V است که به ازاء آن جریان مایع به خارج راکتور قطع می شود. در نهایت ما از روابط ساده زیر برای بیان آنتالپیها استفاده می کنیم (اندیسهای J و ρ به ترتیب متعلق به مایع اصلی فرایند و مایع خنک کننده می باشند)

$$h_J = C_J T_J \quad ; \quad h_\rho = C_\rho T \quad (15-5)$$

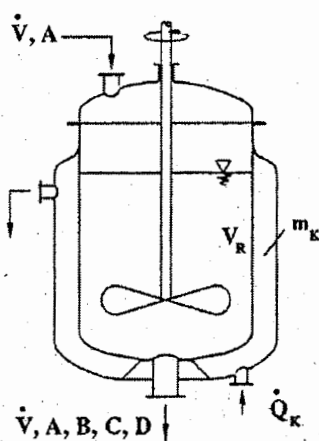
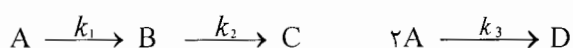
اگر ناحیه انتقال حرارت با تغییر حجم مایع درون راکتور تغییر کند آنگاه حجم مایع درون راکتور نیز باید بعنوان یک متغیر در نظر گرفته شود و معادله بیانگر ارتباط بین مساحت ناحیه منتقل کننده حرارت و حجم مایع درون راکتور باید به دستگاه معادلات افزوده شود. اگر راکتور بصورت یک استوانه عمودی با کف مسطحی به قطر D باشد و ژاکت خنک کننده قسمت کف راکتور را دربرنگرفته باشد آنگاه داریم:

$$A_H = 4 \frac{V}{D} \quad (16-5)$$

فرض کرده‌ایم که ضریب کلی انتقال حرارت U ثابت هستند. در حالت کلی این ضریب می‌تواند تابعی از F_r (سرعت جریان آب خنک‌کننده) باشد که در اینصورت یک متغیر و یک معادله به دستگاه معادلات افزوده خواهد شد.

۵-۲-۴- توصیف CSTR تحت کنترل

سیستم مورد نظر یک راکتور تانک همزن پیوسته با ژاکت خنک‌کننده است که مدلسازی تئوری فرایند بر اساس خواص فیزیکی فرایند واقعی انجام شده است. شکل (۵-۱) دیاگرام شماتیک راکتور را نشان می‌دهد. واکنش اصلی این راکتور تبدیل سیکلوپنتادین^۱ (ماده A) به سیکلوپنتنل (ماده B) بکمک کاتالیزور اسید در محیط آبی می‌باشد. علاوه بر واکنش اصلی، سیکلوپنتادین طی یک واکنش ناخواسته موازی به دی‌سیکلوپنتادین (ماده D) تبدیل می‌شود. همچنین سیکلوپنتادیول (ماده C) نیز طی یک واکنش ناخواسته متوالی از سیکلوپنتنل بدست می‌آید. این واکنش که بنام واکنش واندر ووس^۲ معروف می‌باشد بصورت زیر نمایش داده می‌شود:



شکل ۵-۱- دیاگرام شماتیک سیستم CSTR

^۱ Cyclopentadiene
^۲ Van der vuuse

جریان تغذیه که با نرخ \dot{V} وارد راکتور می‌شود تنها شامل سیکلوپنتادین (ماده A) با غلظت c_{A0} و درجه حرارت T_0 می‌باشد. گرمای واکنش توسط یک مبدل حرارتی خارجی با نرخ \dot{Q}_k از خنک‌کننده بخارج منتقل می‌شود. دینامیک راکتور دارای درجه غیرخطی بالا بوده و بوسیله معادلات دیفرانسیل زیر بیان می‌شود که این معادلات از روابط تعادل اجزاء برای مواد A و B و همچنین تعادل انرژی مابین راکتور و ژاکت خنک‌کننده بدست آمده‌اند [۲۹].

• تعادل اجزاء

$$\dot{c}_A = \frac{\dot{V}}{V_R} (c_{A0} - c_A) - k_1(T) c_A - k_3(T) c_A^2 \quad (17-5)$$

$$\dot{c}_B = -\frac{\dot{V}}{V_R} c_B + k_1(T) c_A - k_2(T) c_B \quad (18-5)$$

• تعادل انرژی

$$\dot{T} = -\frac{\dot{V}}{V_R} (T_0 - T) - \frac{1}{\rho C_P} \left[k_1(T) c_A \Delta H_{RAB} + k_2(T) c_B \Delta H_{RBC} + k_3(T) c_A^2 \Delta H_{RAD} \right] + \frac{k_w A_R}{\rho C_P V_R} (T_J - T) \quad (19-5)$$

$$\dot{T}_J = \frac{1}{m_J C_{PK}} \left[\dot{Q}_K + k_w A_R (T - T_J) \right] \quad (20-5)$$

غلظت مواد A و B بترتیب c_A و c_B می‌باشند. همچنین درجه حرارت‌های راکتور و ژاکت خنک-کننده بترتیب با T و T_J نمایش داده شده‌اند. فرض شده است که سرعت‌های واکنش k_i مطابق قانون آرنیوس بستگی به درجه حرارت داشته باشند:

$$k_i(T) = k_{i0} \cdot \exp\left(\frac{E_i}{R(T + 273.15)}\right), \quad i=1,2,3 \quad (21-5)$$

متغیر خروجی درجه حرارت راکتور است. نرخ جریان ماده A در ورودی راکتور \dot{V} و نرخ تبادل حرارت صورت گرفته از طریق ژاکت خنک کننده \dot{Q}_k نیز متغیرهای ورودی سیستم هستند. برخی از ویژگیهای قابل توجه سیستم CSTR از این قرار می‌باشند:

- ۱- علامت بهره حالت ماندگار در نقطه کار تغییر می‌کند. بنابراین کنترلرهای خطی (با عمل انتگرالی) نمی‌توانند این راکتور را پایدار کرده و عملکرد آنرا بحالت مطلوب برسانند.
- ۲- خاصیت پایداری دینامیک صفر در نقطه کار تغییر می‌کند. بنابراین رفتار کیفی سیستم CSTR در نقاط کار متفاوت و اغتشاشهای مختلف تغییرات قابل توجهی خواهد داشت.
- ۳- سیستم CSTR دارای کاربردهای فراوانی در فرایندهای شیمیائی می‌باشد.

بنا بدلائل اقتصادی عملکرد راکتور در نقطه بهینه بسیار مورد علاقه ما می‌باشد اما رفتار راکتور در این نقطه عملکرد بهینه شدیداً غیرخطی با دینامیک صفر ناپایدار می‌باشد که کنترل راکتور را دشوار می‌سازد. از آنجا که در نقطه کار بهینه مشکلاتی مانند تغییر بهره حالت ماندگار وجود دارند لذا این مساله کنترل اصولاً چالش‌هایی بهمراه دارد. جدول (۵-۱) مقادیر پارامترهای سیستم را نشان می‌دهد. [۲۹]

سیستم CSTR بعلت خواص غیرخطی مثل کندی پاسخ سیستم و سایر مشکلاتی که در مسیر کنترل آن وجود دارند بعنوان یک سیستم معیار برای تست عملکرد انواع طرحهای کنترلی بکار رفته است اما اکثریت قریب به اتفاق این روشها از مدلهای ساده CSTR استفاده کرده‌اند. در این مدلها فرض می‌شود درجه حرارت ژاکت خنک‌کننده بطور مستقیم و بسرعت قابل تغییر است بنابراین از دینامیک راکتور صرفنظر شده و معادله (۵-۲۰) حذف می‌شود. تقریب در نظر گرفته شده تنها هنگامی معتبر است که دینامیک ژاکت در مقایسه با دینامیک راکتور خیلی سریع باشد که با توجه به جرم راکتور و تاخیرهای زمانی موجود، چنین فرضی غالباً باعث پائین آمدن دقت مدل می‌شود همچنین در این مدلها از تغییرات C_B صرفنظر می‌شود بنابراین معادله (۵-۱۸) نیز حذف می‌شود و دو معادله باقیمانده نیز ساده‌تر می‌شوند. بنابراین تا حدود زیادی از دشواریهای کنترل راکتور کاسته می‌شود. در این پایان‌نامه از مدل غیرخطی کامل راکتور استفاده شده است و بهنگام اعمال کنترلر gpc فازی هیچگونه ساده‌سازی یا خطی-

سازی در معادلات سیستم انجام نشده است. همچنین برای استفاده از کنترلر gpc خطی، معادلات مدل کامل بصورت زیر خطی سازی شده اند:

$$A = \begin{bmatrix} \frac{\dot{V}}{V_R} - k_1 - 2k_3c_A & 0 & -\frac{\partial k_1}{\partial T}c_A - \frac{\partial k_1}{\partial T}c_A^2 & 0 \\ k_1 & -\frac{\dot{V}}{V_R} - k_2 & -\frac{\partial k_1}{\partial T}c_A - \frac{\partial k_2}{\partial T}c_B & 0 \\ A(3,1) & -\frac{1}{\rho C_P}k_2\Delta H_{RBC} & A(3,3) & \frac{k_w A_R}{\rho C_{PJ}V_R} \\ 0 & 0 & \frac{k_w A_R}{m_J C_{PJ}} & -\frac{k_w A_R}{m_J C_{PJ}} \end{bmatrix}$$

$$A(3,1) = -\frac{1}{\rho C_P} (k_1 \Delta H_{RAB} + 2k_3c_A \Delta H_{RAD})$$

$$A(3,3) = -\frac{\dot{V}}{V_R} - \frac{1}{\rho C_P} \left(\frac{\partial k_1}{\partial T}c_A \Delta H_{RAB} + \frac{\partial k_2}{\partial T}c_B \Delta H_{RBC} + \frac{\partial k_3}{\partial T}c_A^2 \Delta H_{RAD} \right) - \frac{k_w A_R}{\rho C_{PJ}V_R}$$

$$B = \begin{bmatrix} c_{A0} - c_A & 0 \\ -c_B & 0 \\ T_0 - T & 0 \\ 0 & \frac{1}{m_K C_{PK}} \end{bmatrix}$$

$$\frac{\partial k_i}{\partial T} = \frac{k_i(T) E_i}{R(T+273.15)^2}, \quad i = 1, 2, 3$$

نام پارامتر	علامت	مقدار پارامتر
ضریب برخورد ^۱ واکنش k_1	k_{10}	$۱/۲۸۷ \times ۱۰^{۱۲} \text{ h}^{-۱}$
ضریب برخورد واکنش k_2	k_{20}	$۱/۲۸۷ \times ۱۰^{۱۲} \text{ h}^{-۱}$
ضریب برخورد واکنش k_3	k_{30}	$۹/۰۴۳ \times ۱۰^۹ \text{ h}^{-۱}$
انرژی فعالسازی واکنش k_1	E_1	$-۹۷۸۵/۳ \text{ K}$
انرژی فعالسازی واکنش k_2	E_2	$-۹۷۸۵/۳ \text{ K}$
انرژی فعالسازی واکنش k_3	E_3	-۸۵۶۰ K
سطح ژاکت خنک کننده	A_R	$۰/۲۱۵ \text{ m}^2$
حجم راکتور	V_R	$۰/۰۱ \text{ m}^3$
جرم خنک کننده	m_J	۵ kg
آنتالپی واکنش k_1	H_{AB}	$۴/۲ \frac{\text{kJ}}{\text{molA}}$
آنتالپی واکنش k_2	H_{BC}	$-۱۱ \frac{\text{kJ}}{\text{molB}}$
آنتالپی واکنش k_3	H_{AD}	$-۴۱/۸۵ \frac{\text{kJ}}{\text{molA}}$
چگالی	ρ	$۰/۹۳۴۲ \frac{\text{kg}}{\text{L}}$
ظرفیت حرارتی	C_P	$۳/۰۱ \frac{\text{kJ}}{\text{kg.K}}$
ضریب انتقال حرارتی برای ژاکت خنک کننده	k_w	$۴۰۳۲ \frac{\text{kJ}}{\text{h.m}^2.\text{K}}$
ظرفیت حرارتی خنک کننده	C_{PJ}	$۲ \frac{\text{kJ}}{\text{mol.K}}$

جدول ۵-۱- مقادیر پارامترهای سیستم CSTR

^۱. Collision Factor

۵-۳- شبیه‌سازی

در این قسمت با بکار بردن کنترلر فازی GPC در کنترل یک سیستم غیرخطی معیار (CSTR) تلاش می‌کنیم که بعضی از تواناییهای این طرح در شناسائی و کنترل همزمان سیستمهای غیرخطی ناشناخته را به نمایش بگذاریم. برای کنترل CSTR با کنترلرهای پیش‌بین به تمام متغیرهای حالت سیستم نیاز داریم که شامل درجه حرارت‌های راکتور و ماده خنک‌کننده و غلظت‌های مواد A و B در راکتور می‌باشند. اندازه‌گیری درجه حرارت‌های مواد درون راکتور و ماده خنک‌کننده آسان است اما اندازه‌گیری غلظت مواد کار مشکلی می‌باشد بنابراین در تمام شبیه‌سازها فرض می‌کنیم غلظت‌های مواد A و B قابل اندازه‌گیری نمی‌باشند و برای تخمین مقادیر ایندو متغیر از فیلتر کالمن استفاده می‌کنیم. ترکیب کنترلر-های پیش‌بین با رویکرد حالت در بخش ۲-۶-۳ مورد بحث قرار گرفته است.

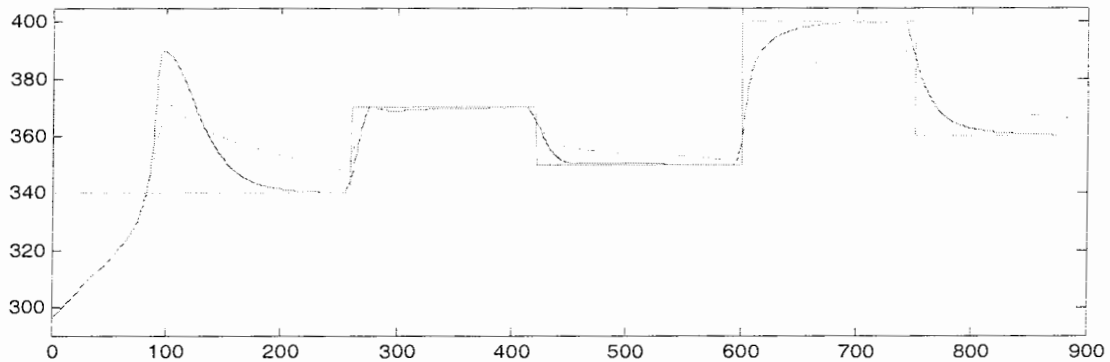
۵-۳-۱- ردیابی سیگنال مرجع

در فصل چهار دو نوع کنترلر مورد بحث قرار گرفت. یکی کنترلر فازی-تطبیقی که از شناسائی برخط مدل فازی-تطبیقی TS و خطی‌سازی آن استفاده می‌کند و دیگری کنترلر GPC خطی که بجای یک مدل پیش‌بین استاتیک از خطی‌سازی ژاکوبین مغادلات غیرخطی سیستم CSTR استفاده می‌کند. در این قسمت عملکرد دو کنترلر مذکور با یکدیگر مقایسه می‌شود.

در ابتدا برای بررسی حالت راه‌اندازی سیستم فرض می‌کنیم دمای اولیه راکتور و ژاکت هر دو مساوی دمای ماده A در ورودی راکتور هستند و راکتور حاوی هیچکدام از مواد A یا B نمی‌باشد:

$$T = T_J = T_0 \quad ; \quad c_A = c_B = 0$$

شکل ۵-۲ عملکرد کنترلرهای GPC فازی و خطی در ردیابی سیگنال مرجع با شرایط اولیه فوق را نشان می‌دهد.



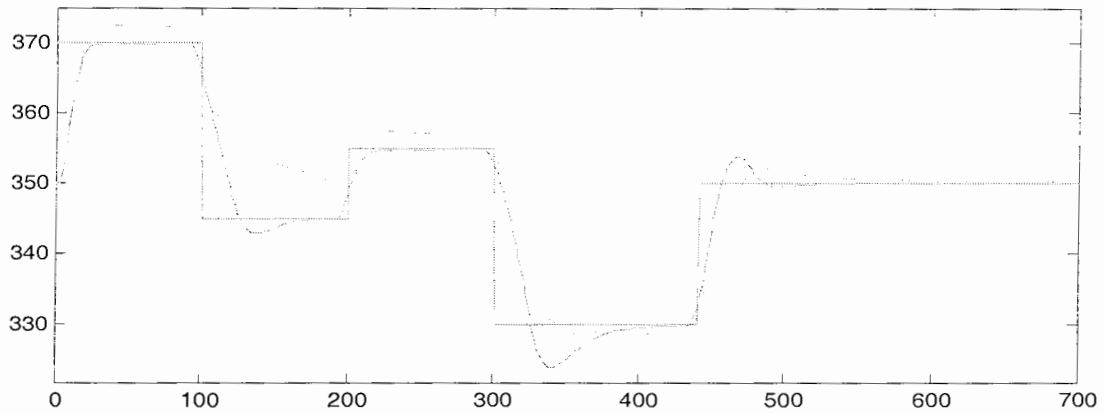
شکل ۵-۲- ردیابی منحنی مرجع پس از راه‌اندازی { پاسخ کنترلر فازی تطبیقی (—) ، پاسخ کنترلر خطی (---) }

ملاحظه می‌شود که پاسخ سیستم در ابتدا مقداری فرارفت^۱ دارد. علت اینست که سیستم CSTR بخاطر خواص غیرخطی در نقاط کار مختلف دارای رفتارهای متفاوتی می‌باشد و شرایط اولیه در نظر گرفته شده برای راه‌اندازی سیستم فاصله زیادی با شرایط عملکرد بهینه دارند بنابراین راه‌اندازی CSTR و هدایت پارامترهای آن از شرایط اولیه به شرایط بهینه با دشواریهای بیشتری همراه می‌باشد. گذشته از این فرایند شیمیائی انجام شده گرمازا می‌باشد بنابراین راکتور بطور ذاتی تمایل به گرم شدن دارد.

بعد از گذشت حالت گذرا مشاهده می‌شود که پاسخ بدست آمده از کنترلر GPC فازی-تطبیقی سرعت به منحنی مرجع همگرا می‌شود اما پاسخ بدست آمده از کنترلر GPC بسیار کند است و نمی‌تواند تغییرات منحنی مرجع را با سرعت مناسبی دنبال کند.

در ادامه فرض می‌کنیم سیستم در لحظات اولیه در نزدیکی نقطه بهینه کار می‌کند. عملکرد کنترلرهای GPC فازی و خطی در شکل ۵-۳ نشان داده شده‌اند. مجدداً می‌بینیم که پاسخ کنترلر فازی-تطبیقی بسیار سریعتر همگرا می‌شود. آشکار است که با دور شدن منحنی مرجع از مقادیر بهینه (حوالی ۳۵۰ درجه) پاسخ کنترلر خطی دارای خطای حالت ماندگار قابل توجهی می‌شود علت این امر تا حدود زیادی به افت دقت مدل خطی ژاکوبین در پیش‌بینی رفتار سیستم مربوط می‌شود در حالیکه مدل خطی بدست آمده از خطی‌سازی مدل TS در تمام شرایط می‌تواند بخوبی رفتار آینده سیستم را تخمین بزند.

^۱ Overshoot

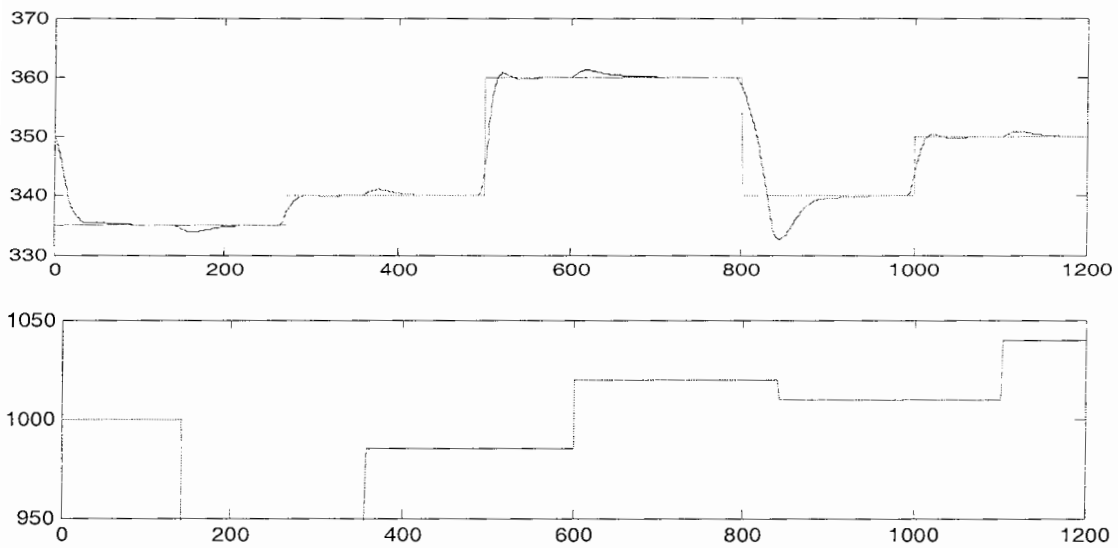


شکل ۳-۵- ردیابی منحنی مرجع در شرایط عادی { پاسخ کنترلر فازی تطبیقي (---) ، پاسخ کنترلر خطی (-.-.-) }

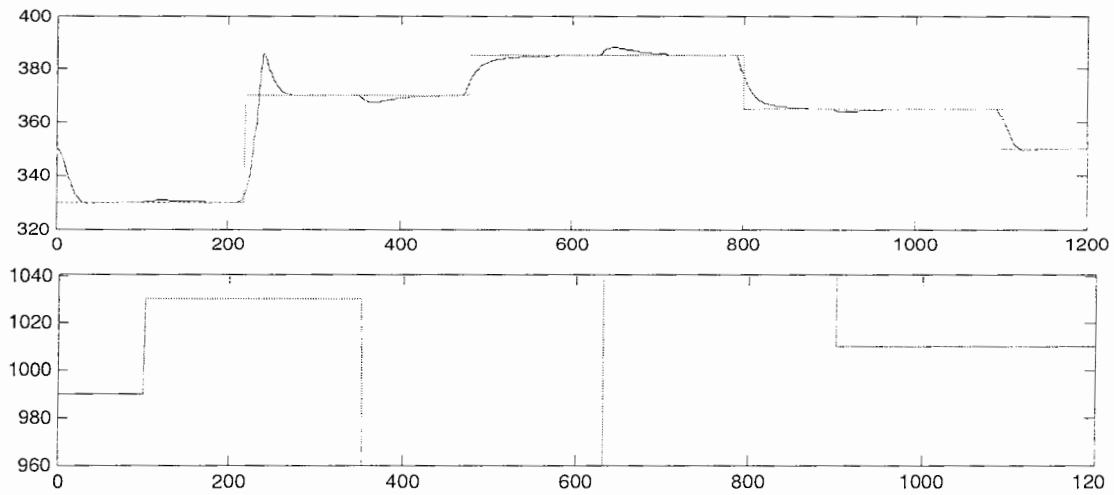
۵-۳-۲- بررسی مقاومت الگوریتم در برابر تغییرات پارامترها

در این قسمت تغییرات مختلفی به پارامترهای قابل تغییر سیستم اعمال کرده و مقاومت کنترلر GPC فازي در برابر این تغییرات را مورد بررسی قرار می‌دهیم.

۵-۳-۲-۱- تغییرات غلظت ماده A در ورودی راکتور (C_{A0})



شکل ۴-۵- تغییرات غلظت ماده A در ورودی راکتور و اثر آن روی خروجی (الف)

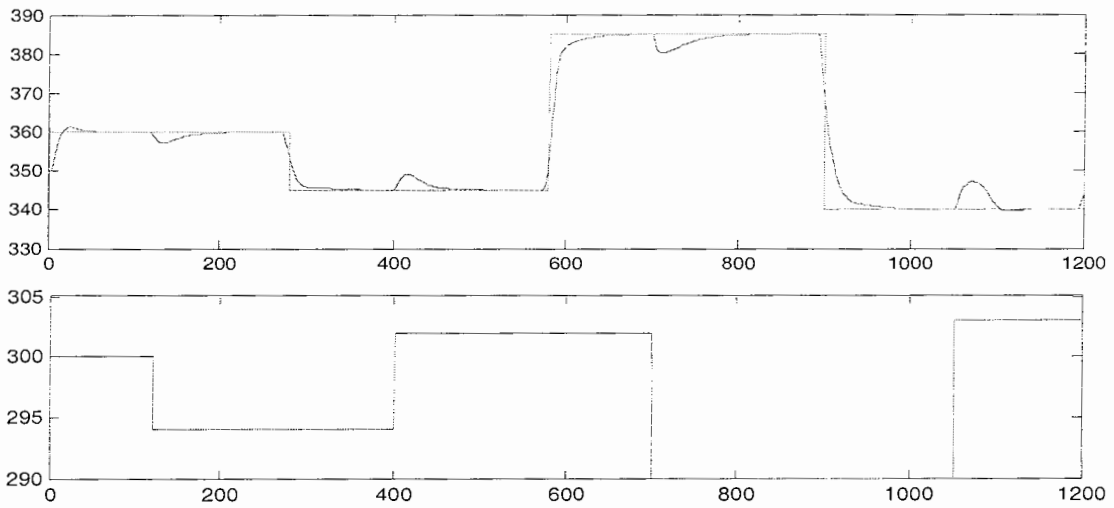


شکل ۴-۵- تغییرات غلظت ماده A در ورودی راکتور و اثر آن روی خروجی (ب)

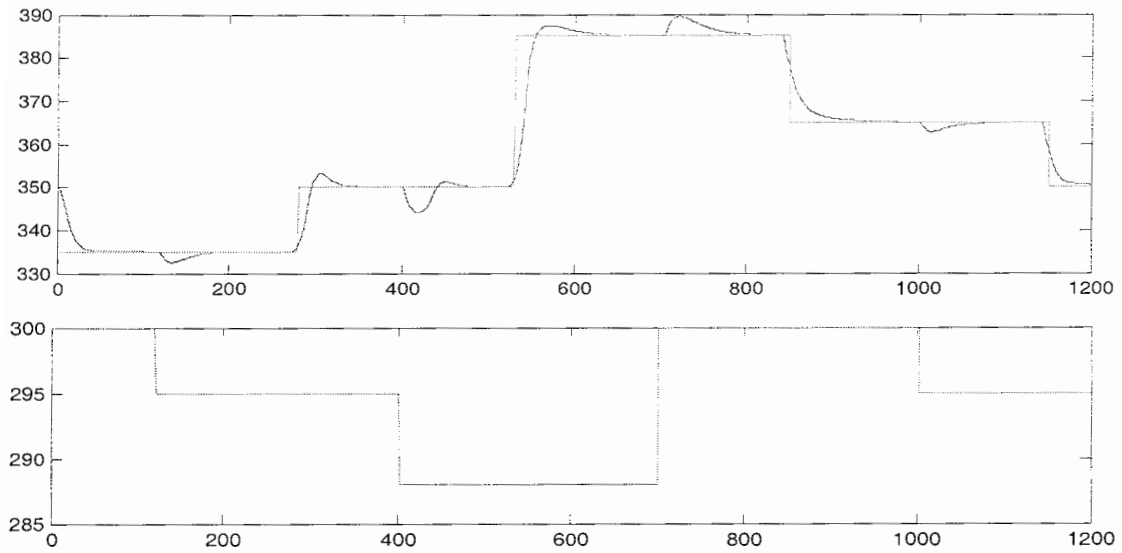
شکل ۴-۵ بوضوح نشان می‌دهد که بعد از هر تغییر در غلظت ماده A در ورودی راکتور، شناساگر-کنترلر فازی-تطبیقی بخوبی می‌تواند تغییر مذکور را شناسائی کند و انحراف خروجی واقعی از خروجی مطلوب را بدون هیچگونه نوسان سرعت میرا کند.

۵-۳-۲- تغییرات دمای ماده A در ورودی سیستم (T_0)

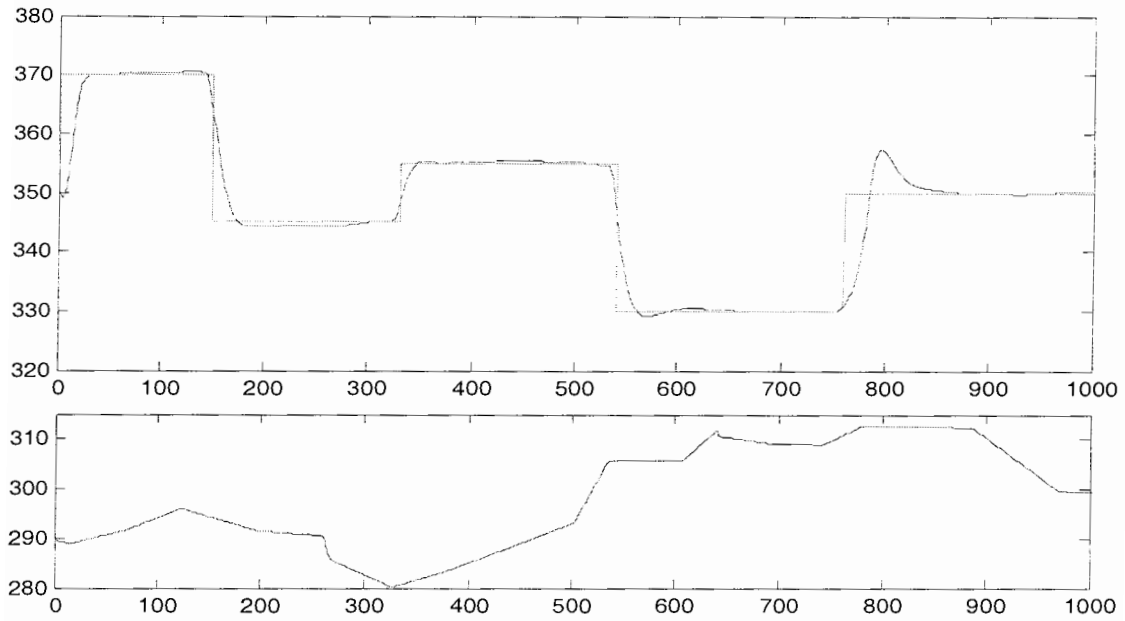
مهمترین اغتشاش وارده به CSTR تغییر دمای ماده A در ورودی سیستم می‌باشد که بطور مستقیم می‌تواند خروجی سیستم را تحت تاثیر قرار دهد. مقایسه دامنه انحراف خروجی واقعی از خروجی مرجع در لحظات تغییر پارامترها در شکل‌های (۵-۵) و (۴-۵) نشان می‌دهد که سیستم به تغییر دمای ماده A بیشتر از تغییر غلظت آن در ورودی حساس می‌باشد. کنترلر فازی در هر دو حالت تغییر ناگهانی و تغییر تدریجی و مداوم دمای ماده A در ورودی هم بخوبی توانسته است انحرافات بوجود آمده را بدون نوسان میرا کرده و عمل ردیابی را ادامه دهد.



شکل ۵-۵- تغییرات دمای ماده A در ورودی راکتور و اثر آن روی خروجی (الف)

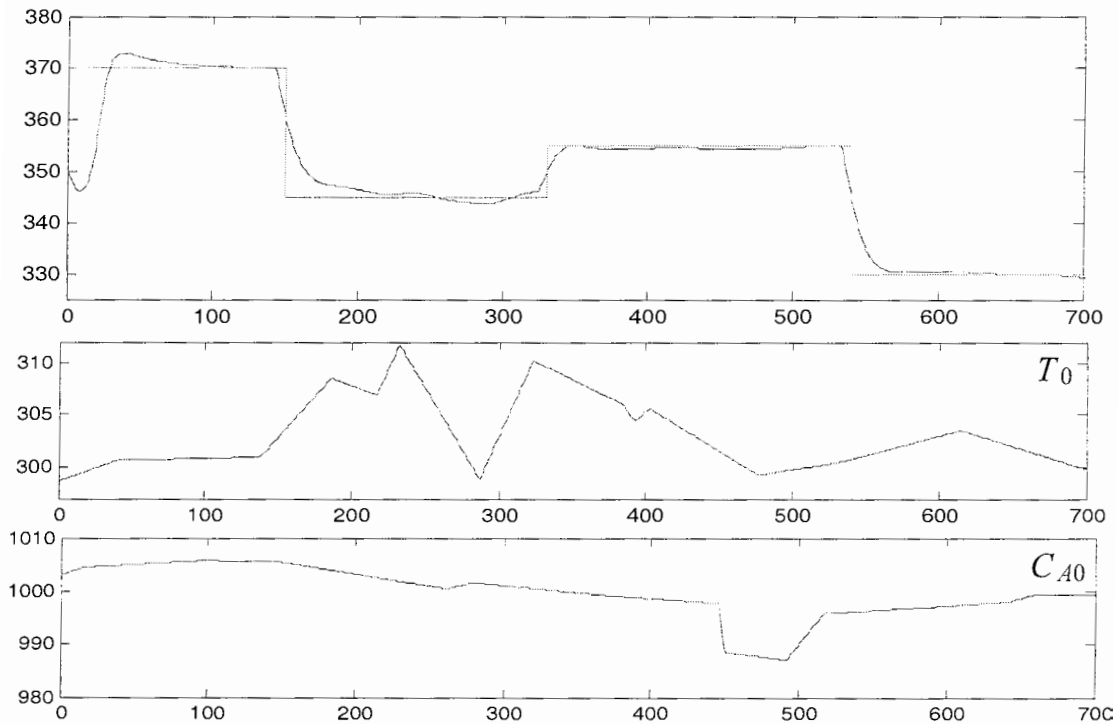


شکل ۵-۵- تغییرات دمای ماده A در ورودی راکتور و اثر آن روی خروجی (ب)



شکل ۵-۶- تغییرات تدریجی ، مداوم و تصادفی دمای ماده A در ورودی راکتور و اثر آن روی خروجی

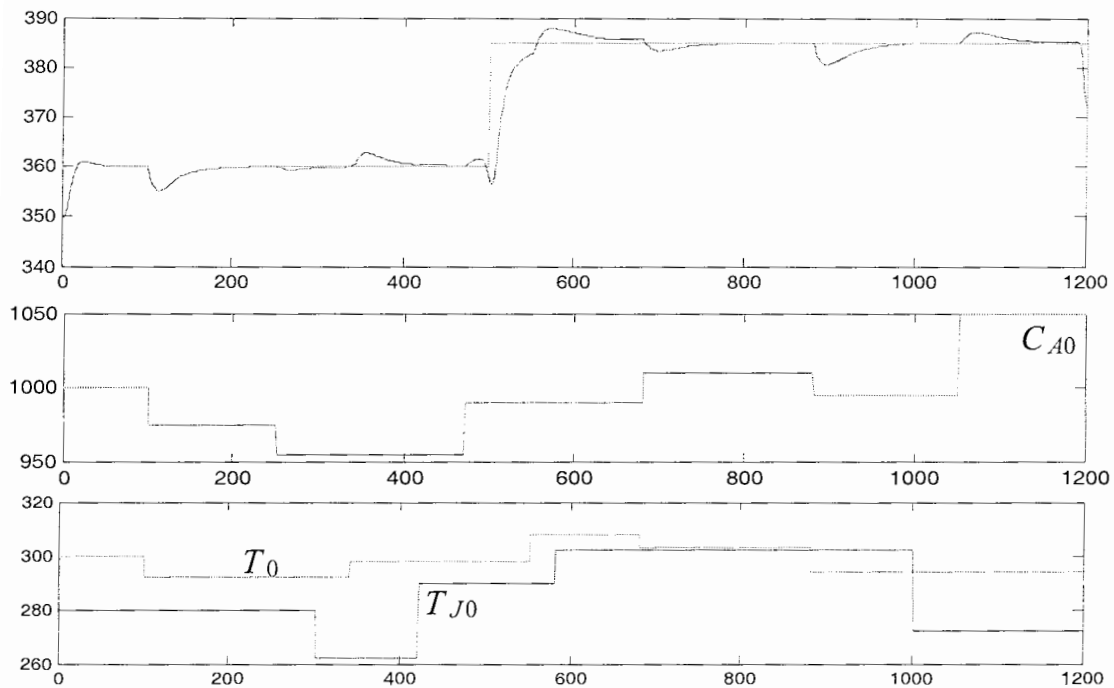
۵-۳-۲-۳- تغییرات همزمان دو پارامتر T_0 و C_{A0}



شکل ۵-۷- تغییرات تدریجی و تصادفی دما و غلظت ماده A در ورودی و اثر آن روی خروجی

۵-۳-۴- تغییرات همزمان سه پارامتر T_{J0} و T_0 ، C_{A0}

در این بخش سعی شده است بدترین شرایط ممکن از نظر تغییر پارامترها برای سیستم ایجاد شود باین شکل که اضافه بر تغییرات تصادفی در دما و غلظت ماده A در ورودی سیستم تغییراتی نیز در دمای ماده خنک‌کننده در ورودی ژاکت اعمال می‌شود.



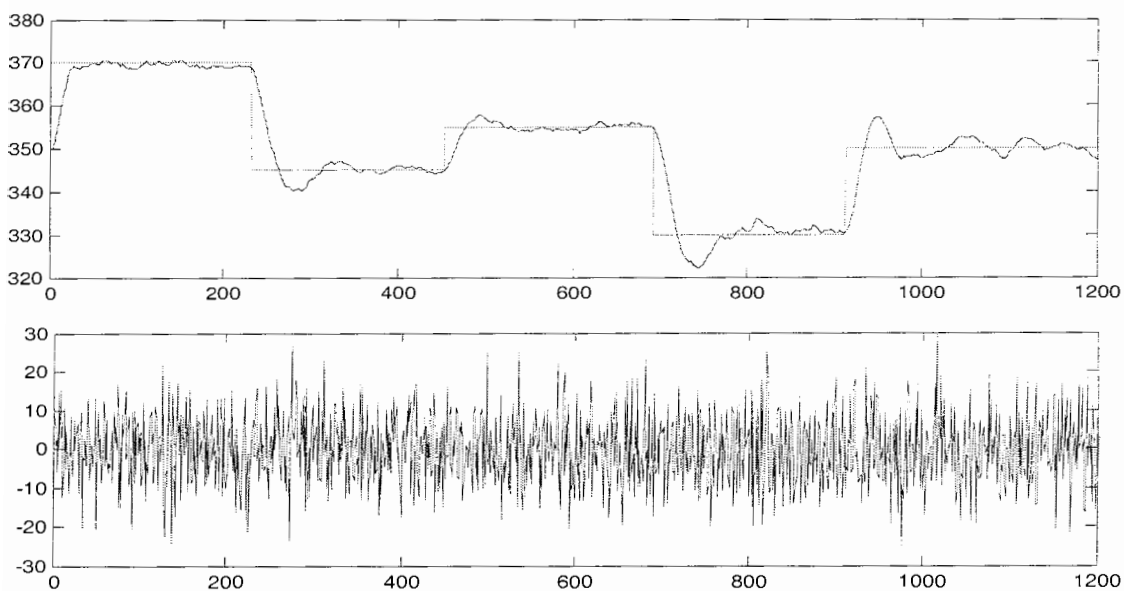
شکل ۵-۸- تغییرات همزمان سه پارامتر و اثر آن در ردیابی سیگنال مرجع

از شکل ۵-۸ دیده می‌شود که حتی تغییر پارامترها در حالات گذرا نیز نمی‌تواند سیستم را ناپایدار کرده یا عمل ردیابی را با مشکلی مواجه کند. بعنوان مثال در لحظه $t=500$ سیگنال مرجع از ۳۶۰ به ۳۸۵ تغییر کرده است و سیستم بعد از $t=500$ در حالت گذرای ناشی از این تغییر می‌باشد. در لحظه $t=550$ و قبل از اتمام حالت گذرا پارامتر T_0 باندازه ۱۰ درجه افزایش پیدا کرده است. در این لحظات سیستم بعد از یک فراجهدش کوچک دوباره به منحنی مرجع همگرا شده است. در حدود $t=580$ پارامتر T_{J0} نیز

تغییر کرده است. قبل از اتمام حالت گذرای مربوطه به دو تغییر مذکور، دو پارامتر T_0 و C_{A0} تغییر کرده‌اند که باز هم انحراف خروجی سرعت و بدون نوسان میرا شده است.

۵-۳-۳- بررسی اثر وجود نویز در اندازه‌گیری خروجی

در این قسمت نویز گوسی سفید را بعنوان نویز اندازه‌گیری با خروجی جمع کرده و اثر آن بر ردیابی سیگنال مرجع را بررسی می‌کنیم. نویز اندازه‌گیری و خروجی سیستم در شکل (۵-۹) نشان داده شده‌اند. ملاحظه می‌شود که خروجی واقعی سیستم باز هم بخوبی توانسته است سیگنال مرجع را ردیابی کند. از مقایسه نویز اندازه‌گیری با خروجی نتیجه می‌شود که دامنه تغییرات تصادفی موجود در خروجی واقعی بمیزان قابل توجهی از دامنه نویز اندازه‌گیری کوچکتر می‌باشد که این خاصیت حذف نویز تنها یکی از مزایای استفاده از شناساگر-کنترلر فازی است.



شکل ۵-۹- نویز جمع‌شونده با خروجی در حین اندازه‌گیری و اثر آن بر ردیابی سیگنال مرجع

نتیجه‌گیری و پیشنهاد

شبیه‌سازی‌های مختلف انجام شده در فصل چهار نشان دادند که شناساگر فازی-عصبی می‌تواند سیستم‌های دینامیکی، غیرخطی و متغیر با زمان را با دقت بالایی شناسایی کند و مدل‌های فازی بدست‌آمده دارای ساختار فشرده و حجم محاسبات قابل قبولی می‌باشند. همچنین از شبیه‌سازی‌های فصل پنج نتیجه می‌شود کنترلر فازی-تطبیقی که از ترکیب الگوریتم شناسایی برخط مدل فازی TS، خطی‌سازی مدل مذکور و الگوریتم GPC خطی بدست آمده است در کنترل سیستم غیرخطی معیار CSTR در شرایط کاری مختلف و در حضور نویز و تغییر پارامترها کاملاً موفق بوده است. همچنین مقایسه عملکرد کنترلر GPC فازی-تطبیقی با کنترلر GPC خطی که از خطی‌سازی ژاکوبین استفاده می‌کند برتری روش GPC فازی-تطبیقی را بوضوح نشان می‌دهد.

بعلت اینکه در روش بهینه‌سازی $B \& B$ جستجوی انجام شده در فضاها محدود بصورت تصادفی انجام می‌شود پیشنهاد می‌شود که از روش بهینه‌سازی حرکت اجتماع ذرات (PSO)^۱ استفاده شود. روش مذکور مبتنی بر الگوهای مورد استفاده توسط پرندگان برای یافتن بهترین مسیر حرکت می‌باشد و در اصل یک روش بهینه‌سازی خارج‌خط است که برای کاربرد موثر در الگوریتم $B \& B$ باید اصلاحاتی روی آن انجام شود. در این الگوریتم جمعیت اولیه پاسخها بصورت تصادفی انتخاب می‌شود که در این مرحله می‌توانیم قیود موجود را اعمال کنیم باین‌شکل که جمعیت اولیه باید در فضای مجاز پراکنده شده باشد. همچنین در حین اجرای الگوریتم و با حرکت نقاط بسمت جواب بهینه اگر نقطه‌ای از مرزهای تعیین شده بوسیله قیود مذکور خارج شد باید به ناحیه مجاز برگردانده شود.

^۱ Particle Swarm Optimization

مراجع

- [1] Ruiyun Qi and Mietek A. Brdys, “ Adaptive Fuzzy Modeling and Control for Discrete Time Nonlinear Uncertain Systems”, 2005 American Control Conference , June 8-10 , 2005 . Portland, OR, USA
- [2] N. K. Kasabov and Q.Song , “DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and its Application For Time-Series Prediction” IEEE Trans. Fuzzy Syst., vol. 10, pp. 144-154 , Apr. 2002
- [3] Plamen Angelov and Xiao-Wei Zhou , “Evolving Fuzzy Systems from Data Streams in Real-Time” , IEEE Transaction on Fuzzy Systems.,vol.12 , NO. 1 , pp.365-378 , Februray 2004
- [4] Plamen Angelov and Dimitar P.Filev, “An Approach to Online Identification of Takagi-Sugeno Fuzzy Models” , IEEE Trans.Syst.,Man Cybern.,vol.34 , NO. 1 , pp.484-498 , Februray 2004
- [5] Plamen Angelov , Costas Xydeas and Dimitar Filev , “On-line Identification of MIMO Evolving Takagi-Sugeno Fuzzy Models”, IEEE Conf. pp. 55-60 , July 2004 , Budapest. Hun.
- [6] Zhengrong Li and Meng Joo Er , “A Nonlinear Transversal Fuzzy Filter With Online Clustering”, pp.1585-1593 , 5th Asian Control Conference ,2004

- [7] Kian Hong Quah and Chai Quek , “Online Local Learning With Generic Fuzzy Input Takagi-sugeno-kang Fuzzy Framework for Nonlinear System Estimation ”, IEEE Trans.Syst.,Man Cybern.,vol.36 , no. 1, February 2006
- [8] D.P.Filev. “Rule-base guided adaptation for mode detection in process control. ” in proc. Joint 9th IFSA World Congr/20th NAFIPS Annu. Conf. Vancouver,BC,Canada, July 2001. pp. 1068-1073
- [9] D.P.Filev, T.Larson, and L.Ma, “Intelligent control for automotive manufacturing—rule based guided adaptation” , in proc.IEEE Conf .IECON’00 , Nagoya , japan , oct. 2000 , pp .283-288
- [10] T.Takagi and M.Sugeno ,“Fuzzy identification of systems and its application to modeling and control” IEEE Trans.Syst.,Man Cybern.,vol.15,pp.116-132,1985
- [11] M.Sugeno and M.Yasukawa, “ A fuzzy logic based approach to qualitative modeling ” IEEE Trans. Fuzzy Syst., vol.1 , pp. 7-31 ,1993
- [12] B.Kosko, “Fuzzy systems as universal approximator”, IEEE Trans.computers., vol .43 , pp 1329-1333 , Nov .1994
- [13] R.Rovatti , “Fuzzy piecewise multilinear and piecewise linear systems as universal approximators in Sobolov norms”, IEEE Trans. Fuzzy Syst., vol 6 , pp 235-249 , May 1998

- [14] S.Barada , and H.Singh , “Generating optimal Adaptive fuzzy neural models for dynamical systems with application to control ”, IEEE Trans.Syst.,Man Cybern.,vol. 28 , no. 3 , pp 371-391 , Aug. 1998 .
- [15] J.Abonyi, R.Babuska, F.Szeifert, “Modified Gath-Geva fuzzy clustering for identification of T-S fuzzy models ” , IEEE Trans.Syst.,Man Cybern.,vol. 32 , no. 5 , pp 612-621 , Oct. 2002
- [16] F.-J.lin , C.-H. Lin , and P.-H Shen, “Self-constructing Fuzzy Neural Network Speed Controller for Permanent-Magnet Synchronous Motor Drive ” , IEEE Trans. Fuzzy Syst., vol. 9 , pp. 751-759 , Oct . 2001 .
- [17] Debrup Chakraborty and Nikhil R.Pal , “A Neuro-Fuzzy Scheme for Simultaneous Feature Selection and Fuzzy Rule-Based Classification ” , IEEE Trans. ,Neural NET. vol. 15 , pp. 110-122 , Januray 2004
- [18] Mohammad Fazle Azeem , Madasu Hanmandlu , “ Structure Identification of Generalized Adaptive Neuro-Fuzzy Inference Systems ” , IEEE Trans. Fuzzy Syst., vol 11 , no.5 , pp. 666-680 , October 2003
- [19] Shiqian Wu, Meng Joo Er and Yang Gao , “ A Fast Approach for Automatic Generation of Fuzzy Rules by Generalized Dynamic Fuzzy Neural Networks” , IEEE Trans. Fuzzy Syst., vol . 9 , no . 4, pp. 578-594 , August 2001

- [20] S.L.Chiu , “Fuzzy Model Identification Based on Cluster Estimation ” , J.Intel . Fuzzy Syst., vol. 2 , pp. 267-278 , 1994
- [21] C.F. Juang and X.T.Lin , “ A Recurrent Self-Organizing Neural Fuzzy Inference Network ” , IEEE Trans. On Neural Networks, vol. 10 , pp. 828-845 , 1999
- [22] K.B.Cho and B.H.Wang , “ Radial Basis Function Based Adaptive Fuzzy Systems And Their Applications To System Identification And Prediction” , Fuzzy Sets And Systems , vol. 83 , pp. 325-339 ,1996 .
- [23] S.A.Bilings and C.F.Fung , “ Recurrent Radial Basis Function Networks For Adaptive Noise Cancellation ” , Neural Networks , vol. 8 , pp. 207-224 , 1991
- [24] L.Yingwei , N.Sundararajan and P.Saratchandran , “ A Sequential Learning Scheme For Function Approximation By Using Minimal Radial Basis Function Neural Networks ” , Neural Computa. , vol. 9 , pp. 461-478 , 1997.
- [25] S. Wu and M.J.Er , “ Dynamic Fuzzy Neural Networks- a Novel Approach to Function Approximation” , IEEE Trans.Syst.,Man Cybern., vol. 30 , pp. 385-364 , 2000
- [26] J.S.R.Jang , “Anfis : Adaptive Network-based Fuzzy Inference Systems” , IEEE Trans.Syst.,Man Cybern.,vol. 23 , pp. 665-685 , 1993

- [27] Ginovanna Castellano , “ A Neurofuzzy Methodology for Predictive Modeling ” ,
Phd Thesis , University of Bari , 2000
- [28] H.Sunan , T.K.Kiang and L.T.Heng , “ Applied Predictive Control – Advances in
Industrial Control ” , Springer-Verlag , 2002
- [29] Andrey Alexandrovich Tyagunov , “ High Performance Model Predictive Control
For Process Industry” , PHD THESIS , Eindhoven University of Technology
- [30] Victor , J. and A.Dourado, “ Eolving Takagi-Sugeno Fuzzy Models ” , Technical
Report , CISUC
- [31] E.F.Camacho and C.Bordons , “ Model Predictive Control ” , Second Edition ,
Springer-Verlag , 2004
- [32] L.H.Li , H-YE SU, J.Chu and X-P.Guan , “ Novel Fuzzy Predictive Pid Control
For a Class of Nonlinear Systems ” , First International Conference on Machine
Learning and Cybernetics , Beijing , November 2002
- [33] W.Gao , M-L.Zhou , Y-C.Li and T.Zhang , “ An Adaptive Generalized Predictive
Control of Time-Varying Delay System ” , Third International Conference on
Machine Learning and Cybernetics , Shanghai , August 2004

- [34] P.K.Roy , G.K.Mann and C.Hawllader , “ Fuzzy Rule-Adaptive Model Predictive Control For a Multivariable Heating System ” , IEEE Conference on Control Applications , Toronto , Canada , August 2005
- [35] Y.L.Huang , Helen H.Lou , J.P.Gong and Thomas F.Edgar , “ Fuzzy Model Predictive Control ” , IEEE Trans. On Fuzzy Syst. , Vol 8 , No. 6 , December 2000
- [36] B.Wayne Bequette , “ Process Control , Modeling , Design and Simulation ” , Prentice Hall of India , 2004
- [37] C.R .Cutler , and B.L ,Ramaker , 1980 , “ Dynamic Matrix Control – a Computer Control Algorithm ” , Proc Joint Automatic Control Conf , San Francisco , California
- [39] D.W. Clarke, C. Mohtadi, and P. Tuffs , “ Generalized Predictive Control – Parts I and II ” , Automatica , 23: 137-160 , 1987
- [40] M.Mahfouf , D.A. Linkens and M.Abbod , “ Adaptive Fuzzy Tsk Model-Based Predictive Control Using a Carima Model Structure ” , Trans ICHemE , Vol 78 , Part A , May 2000
- [41] B.Liu , Q.Shen , H.Ye su and J.Chu , “ A Nonlinear Predictive Control Algorithm Based on Fuzzy Online Modeling and Discrete Optimization” , IEEE Conference on Control Applications , Toronto , Canada , August 2005

- [42] S.Mollov , R.Babuska , J.Abonyi and H.B.Verbruggen,“ Effective Optimization for Fuzzy Model Predictive Control ” , IEEE Trans. On Fuzzy Systems , Vol 12 , No. 5 , October 2004
- [43] J.A.Roubos , S.Mollov , R.Babuska , H.B.Verbruggen ,“ Fuzzy Model-Based Predictive Control Using Takagi-Sugeno Models ” , International Journal of Approximate Reasoning 22, pp. 3-30 , 1999
- [44] J.M.Sousa , M.Setnes , U.Kaymak ,“ Adaptive Decision Alternatives in Fuzzy Predictive Control Control ” , in : FUZZ-IEEE , Vol. 1 , Anchorage , Alaska , 1998 , pp. 698-703
- [45] L.F.Mendonca , J.M.Sousa , J.M.G. Sa da Costa , “ Optimization Problems in Multivariable Fuzzy Predictive Control ” , International Journal of Approximate Reasoning 36 (2004) , 199-221
- [46] Wen Yu and Xiaoou Li , “ Fuzzy Identification Using Fuzzy Neural Networks With Stable Learning Algorithms ” , IEEE Trans. On Fuzzy Systems , Vol 12 , No.3 , pp. 411-420 , June 2004
- [47] Enso Ikonen , Kaddour Najim , “ Advanced Process Identification and Control ” , Marcell Dekker Inc , New York , 2002

- [48] S.Yuan LI , YI-Peng Yang , “ Online Constrained Predictive Control Algorithm using Multi-objective Fuzzy-Optimization and a Case Study ” , Fuzzy Optimization and Decision Making , 2 , pp. 123-142 , 2003
- [49] L.S.Coelho , R.A.Krohling , “ Predictive Control Tuning using Modified Particle Swarm Optimization Based on Cauchy and Gaussian Distributions ” , Springer-Verlag Berlin Heidelberg 2005

ضمیمه الف

اثبات برخی روابط الگوریتم شناسائی برخط مدل TS

A. الگوریتم بازگشتی وزن دار

الگوریتم بازگشتی وزن دار را می‌توان از شبه‌انعکاس^۱ بدست آورد باینصورت که همانند روش بکار رفته در بدست آوردن RLS از LS، ماتریسها را بصورت مجموع بیان کرده و اجزاء را مرتب کنیم:

$$\hat{\pi}_{ik} = C_{ik} \left(\sum_{j=1}^{k-1} x_{ej} \lambda_1(x_j) y_j + x_{ek} \lambda_1(x_k) y_k \right) \quad (\text{الف-۱})$$

که C_{ik} ماتریس کواریانس می‌باشد:

$$C_{ik} = \left(\sum_{j=1}^k x_{ej}^T \lambda_1(x_j) x_{ej} \right)^{-1} \quad (\text{الف-۲})$$

با استفاده از این عبارت جهت تخمین بر اساس داده $k-1$ و مرتب کردن مجدد خواهیم داشت:

$$\begin{aligned} \sum_{j=1}^{k-1} x_{ej} \lambda_1(x_j) y_j &= C_{ik-1}^{-1} \hat{\pi}_{ik-1} \\ &= C_{ik}^{-1} \hat{\pi}_{ik-1} - x_{ek} \lambda_1(x_k) x_{ek}^T \hat{\pi}_{ik-1} \end{aligned} \quad (\text{الف-۳})$$

جایگذاری (الف-۳) در (الف-۱) و استفاده از قضیه انعکاس ماتریسها عبارت زیر را بدست می‌دهد:

$$\begin{aligned} \pi_{ik} &= \hat{\pi}_{ik-1} - C_{ik} x_{ek} \lambda_1(x_k) x_{ek}^T \hat{\pi}_{ik-1} + C_{ik} x_{ek} \lambda_1(x_k) y_k \\ &= \hat{\pi}_{ik-1} + C_{ik} x_{ek} \lambda_1(x_k) (y_k - x_{ek}^T \hat{\pi}_{ik-1}) \end{aligned} \quad (\text{الف-۴})$$

بنابراین رابطه (۳-۳۵) اثبات می‌شود. با روشی مشابه و از روی تعریف ماتریس کواریانس برای تخمین

بر اساس k داده به عبارت زیر می‌رسیم:

$$\begin{aligned} C_{ik} &= \left(\sum_{j=1}^{k-1} x_{ej}^T \lambda_1(x_j) x_{ej} + x_{ek}^T \lambda_1(x_k) x_{ek} \right)^{-1} \\ &= (C_{k-1}^{-1} + \lambda_1(x_k) x_{ek}^T x_{ek})^{-1} \\ &= C_{ik-1} - C_{ik-1} \lambda_1(x_k) x_{ek} \times (I + \lambda_1(x_k) x_{ek}^T C_{ik-1} x_{ek})^{-1} x_{ek}^T C_{ik-1} \end{aligned} \quad (\text{الف-۵})$$

^۱: Pseudo-inversion

B. محاسبه بازگشتی پتانسیلها

با شروع از رابطه اولیه محاسبه پتانسیل (۳-۱۸) و بیان تصویر فواصل در لحظه زمانی k ام بصورت دقیق خواهیم داشت :

$$p_k(z_k) = \frac{1}{1 + \frac{1}{(k-1)} \sum_{i=1}^{k-1} \sum_{j=1}^{n+1} \left\{ (C_k^j)^2 - 2z_k^j z_i^j + (z_i^j)^2 \right\}} \quad (\text{الف-۶})$$

با مرتب کردن رابطه بالا داریم :

$$p_k(z_k) = \frac{1}{1 + \frac{1}{(k-1)} \left((k-1) \sum_{i=1}^{n+1} (z_i^j)^2 - 2 \sum_{j=1}^{n+1} \sum_{i=1}^{k-1} z_i^j + \sum_{i=1}^{k-1} \sum_{j=1}^{n+1} (z_i^j)^2 \right)} \quad (\text{الف-۷})$$

که معادل رابطه (۳-۱۹) می باشد.

C. بروزرسانی بازگشتی پتانسیلهای نقاط کانونی

اگر داده ای بعنوان نقطه کانونی خوشه-قاعده ای در زمان $k-1$ پذیرفته شود ($c_1 = z_{k-1}$ برای $k > 2$) پتانسیل آن طبق رابطه بصورت زیر محاسبه می شود:

$$p_{k-1}(c_i) = \frac{1}{1 + \frac{1}{(k-2)} \sum_{i=1}^{k-2} \sum_{j=1}^{n+m} (d_{i(k-1)}^j)^2} \quad (\text{الف-۸})$$

مجموع موجود در رابطه بالا را می توان بصورت زیر نوشت :

$$\sum_{i=1}^{k-2} \sum_{j=1}^{n+m} (d_{i(k-1)}^j)^2 = (k-2) \left(\frac{1}{p_{k-1}(c_i)} - 1 \right) \quad (\text{الف-۹})$$

در پله زمانی بعدی (k) جهت در نظر گرفتن اثر داده جدید روی این مرکز ، آنرا بروز می کنیم .

$$p_k(c_i) = \frac{1}{1 + \frac{1}{(k-1)} \left(\sum_{i=1}^{k-2} \sum_{j=1}^{n+m} (d_{i(k-1)}^j)^2 + \sum_{j=1}^{n+m} (d_{k(k-1)}^j)^2 \right)} \quad (\text{الف-۱۰})$$

با جایگذاری (الف-۹) در (الف-۱۰) خواهیم داشت :

$$p_k(c_i) = \frac{1}{1 + \frac{1}{(k-1)} \left((k-2) \left(\frac{1}{p_{k-1}(c_i)} - 1 \right) + \sum_{j=1}^{n+m} (d_{k(k-1)}^j)^2 \right)} \quad (\text{الف-۱۱})$$

با مرتب کردن عبارت فوق به (الف-۱۲) می‌رسیم که معادل رابطه (۳-۲۲) می‌باشد.

$$p_k(c_i) = \frac{(k-1)p_{k-1}(c_i)}{k p_{k-1}(c_i) - p_{k-1}(c_i) + k - 2 - k p_{k-1}(c_i) + 2 p_{k-1}(c_i) + p_{k-1}(c_i) \sum_{j=1}^{n+m} (d_{k(k-1)}^j)^2} \quad (\text{الف-۱۲})$$

D. بروزرسانی ماتریس کواریانس

اجازه دهید از بردار ورودیهای وزن داده شده بوسیله سطوح آتش غیرنرمالیزه قواعد استفاده کنیم:

$$\delta_k = [\varphi_1(x_k) x_{ek}^T, \varphi_2(x_k) x_{ek}^T, \dots, \varphi_R(x_k) x_{ek}^T]^T \quad (\text{الف-۱۲})$$

آنگاه پارامتر زیر را تعریف می‌کنیم.

$$\psi_k = \frac{1}{\sum_{j=1}^R \varphi_j} \delta_k \quad (\text{الف-۱۳})$$

با استفاده از فیلتر کالمن جهت بروزرسانی ماتریسهای کواریانس خواهیم داشت:

$$C_k = C_{k-1} - \frac{C_{k-1} \delta_k \delta_k^T C_{k-1}}{\left(\sum_{j=1}^R \varphi_j \right)^2 + \delta_k^T C_{k-1} \delta_k} \quad (\text{الف-۱۴})$$

و یا با بیان داده‌های موجود تا زمان k بصورت دقیق خواهیم داشت:

$$C_k = \Omega I - \sum_{i=1}^k \frac{A_i}{B_i + F} \quad (\text{الف-۱۵})$$

که در رابطه بالا داریم:

$$A_i = C_{k-1} \delta_k \delta_k^T C_{k-1} \quad ; \quad B_i = \delta_i^T C_{i-1} \delta_i \quad ; \quad F = \left(\sum_{j=1}^R \varphi_j \right)^2 \quad (\text{الف-۱۶})$$

اجازه بدهید فرض کنیم که قاعده اضافه شده در پله زمانی k ام از ابتدا وجود داشته است. آنگاه

ماتریس کواریانس در زمان k باینصورت خواهد بود:

$$\hat{C}_k = \Omega I - \sum_{i=1}^k \frac{A_i}{B_i + \left(\sum_{j=1}^R \varphi_j + \varphi_{R+1} \right)} = \Omega I - \sum_{i=1}^k \frac{A_i}{B_i + F + \delta F_1 + \delta F_1} \quad (\text{الف-۱۷})$$

که داریم $\delta F_1 = \varphi_{R+1}^2$ و $\delta F_2 = 2 \varphi_{R+1} \sum_{j=1}^R \varphi_j$. دیده می‌شود که اضافه کردن قاعده‌ای در پله زمانی k ام باعث انحراف ماتریس کواریانس می‌شود. مقادیر δF_1 و δF_2 خیلی کمتر از ۱ هستند (چون این پارامترها فرمهای مرتبه دوم توابع تعلق می‌باشند) از آنجا که B_i فرم مرتبه دوم حاصلضرب داده ورودی در ماتریس کواریانس است بنابراین می‌تواند رقم بزرگی باشد. F از δF_1 بزرگتر است چون F مجموع N تابع عضویت است اما δF_1 تنها یکی از توابع عضویت می‌باشد. همچنین تنها در صورت برقراری شرط $\varphi_{R+1} > (1/2) \sum_{j=1}^R \varphi_j$ ، F از δF_2 نیز بزرگتر خواهد بود بنابراین اگر ماتریس کواریانس یا همه مقادیر x_{ie} (برای تمام پله‌های زمانی قبلی) به صفر میل کنند آنگاه نقش جمع‌شونده‌ها^۱ مهمتر خواهد بود. آزمایشهای عملی روی تعدادی از توابع نشان می‌دهند که انحراف ماتریس کواریانس با اضافه شدن قاعده جدید، مجانبی خواهد بود.

این تاثیر (غالباً کوچک) را با میانگین معکوس تصحیح نوع میانگین^۲ تقریب می‌زنیم و دلیل منطقی اینکار بدین ترتیب می‌باشد: از روابط (الف-۱۶) و (الف-۱۷)، ماتریس کواریانس انحراف یافته \hat{C}_{k+1} را بصورت تابعی از ماتریس کواریانس اصلی (C_{k+1}) خواهیم داشت:

$$\hat{C}_{k+1} = f(C_{k+1}) \quad (\text{الف-۱۸})$$

بعلت اینکه با افزایش R نقش انحراف ماتریس کواریانس کمتر می‌شود و این وابستگی بفرم توان دوم است یک تقریب از تابع f می‌تواند بصورت معکوس توان دوم میانگین^۳ باشد:

$$\hat{C}_{k+1} = C_{k+1} \frac{R^2 + 1}{R^2} = C_{k+1} \left(1 + \frac{1}{R^2} \right) = \rho C_{k+1} \quad (\text{الف-۱۷})$$

¹ Addends

² Inverse Mean of Average Type of Correction

³ Inverse Squared Mean

ضمیمه ب

چند الگوریتم خوشه‌سازی مهم

A- خوشه‌سازی کروی با استفاده از یادگیری رقابتی بهبود یافته [۲۷]

الگوریتم یادگیری این روش شامل مراحل زیر است :

۱- تعداد خوشه‌ها و مختصات اولیه مراکز آنها باید بر اساس دانش اولیه موجود در مورد توزیع داده‌ها حدس زده می‌شود. مختصات مراکز خوشه‌ها در حقیقت همان مختصات گره‌های لایه سوم شبکه فازی-عصبی است. تعداد این گره‌ها را با R نشان می‌دهیم.

۲- سرعت‌های یادگیری α_w و α_r مقداردهی اولیه می‌شوند بگونه‌ای که $0 \leq \alpha_r \leq \alpha_w \leq 1$

۳- نمونه بعدی داده گرفته می‌شود. $(X(k))$

۴- فاصله بین نمونه داده و تمام مراکز خوشه‌ها محاسبه می‌شود :

$$d(X(k), c_i) = \frac{n_i}{\sum_{j=1}^R n_j} d_E(X(k), c_i) \quad ; i=1:R \quad (1-ب)$$

n_i تعداد دفعاتی است که نورون i ام بعنوان نزدیکترین نورون به داده در حال پردازش شناخته می‌شود و d_E نیز فاصله اقلیدسی می‌باشد.

۵- دو نورون از لایه سوم که از نظر فاصله d به این نمونه نزدیکترند تشخیص داده می‌شوند و بترتیب نورون برنده c_w و نورون رقیب c_r نامیده می‌شوند (مرحله رقابت) بگونه‌ای که داریم :

$$d(X(k), c_w) < d(X(k), c_r) \quad (2-ب)$$

۶- پارامتر n_w بصورت زیر اصلاح می‌شود :

$$n_w = \begin{cases} n_w + 1 & \text{if } r = w \\ n_w & \text{if } r \neq w \end{cases} \quad (\text{ب-۳})$$

۷- مختصات نورونهای برنده و رقیب بگونه‌ای اصلاح می‌شوند که نورون برنده به نمونه داده نزدیکتر شده و نورون رقیب از نمونه داده دورتر شود :

$$c_w = c_w + \alpha_w(X(k) - c_w) \quad (\text{ب-۴})$$

$$c_r = c_r - \alpha_r(X(k) - c_r) \quad (\text{ب-۵})$$

۸- برای تسریع همگرایی الگوریتم ، سرعت‌های یادگیری مطابق روابط زیر کاهش داده می‌شوند (شماره تکرار با k نشان داده شده است)

$$\alpha_w = \alpha_w - \alpha_w * \frac{k}{5} ; \quad \alpha_r = \alpha_r - \alpha_r * \frac{k}{5} \quad (\text{ب-۶})$$

۹- در صورت برقراری شرط زیر به مرحله بعد می‌رویم در غیر اینصورت به مرحله ۳ برمی‌گردیم .

$$\frac{1}{R} \sum_{i=1}^R \|c_i^{k+1} - c_i^k\| \leq \varepsilon \quad (\text{ب-۷})$$

۱۰- نورون‌هایی که مختصات آنها از محدوده تغییرات ورودیها خارج است حذف می‌شوند .

از پارامتر n_i برای حل مشکل گره‌های مرده استفاده شده است . گره‌های مرده هیچگاه برنده نمی‌شوند بنابراین نمی‌توانند بعنوان مرکز خوشه انتخاب شوند. این مشکل عموماً در اثر انتخاب نامناسب موقعیت اولیه این نورونها پیش می‌آید. با توجه به ثابت بودن عدد در نظر گرفته شده برای تعداد خوشه‌ها ، وجود گره‌های مرده ممکن است باعث شود بعضی خوشه‌ها بصورت خوشه‌های جداگانه تشخیص داده نشوند. چون در ابتدای الگوریتم اطلاع دقیقی از نحوه توزیع داده‌ها در دسترس نمی‌باشد نمی‌توانیم این مشکل را در مراحل اولیه حل کنیم . پارامتر n_i باعث می‌شود با هر بار برنده شدن یک نورون ، احتمال برنده شدن آن در تکرار بعد کمتر شود تا گره‌های دیگر نیز شانس برنده شدن داشته باشند . اگر چه در این الگوریتم نیز تعداد خوشه‌ها باید در ابتدا تعیین شود اما با توجه به

مراحل ۷ و حذف نورونهای اضافی در مرحله ۱۰، تعداد خوشه‌ها توسط فرایند یادگیری اصلاح می‌شود بنابراین الگوریتم رقابتی بالا در مقایسه با دیگر روشهای خوشه‌سازی متداول (نظیر میانگین k فازی یا دیگر الگوریتمهای رقابتی) دارای توانائی بالاتری در یافتن تعداد مناسب خوشه‌ها می‌باشد.

B- خوشه‌سازی فوق‌بیضوی^۱ با استفاده از یادگیری رقابتی بهبود یافته [۲۷]

یک خاصیت نامناسب خوشه‌سازی با استفاده از فاصله اقلیدسی اینست که در بعضی موقعیتهای، اهمیت یکسانی به خوشه‌های بزرگ و خوشه‌های کشیده داده می‌شود. همچنین نحوه توزیع داده‌های واقعی بصورت خوشه‌های مجزا یا خوشه‌های گروهی نمی‌باشد بنابراین در خوشه‌سازی فوق‌بیضوی بجای فاصله اقلیدسی از دیگر انواع فاصله نظیر فاصله M^2 استفاده می‌شود. مجذور فاصله M موجود بین نمونه $X(k)$ و بردار مرکز خوشه c_i بصورت زیر محاسبه می‌شود:

$$d_M(X, c_k) = (X - c_i)^T \sum_k^{-1} (X - c_i) \quad (\text{ب-۸})$$

\sum_k^{-1} معکوس ماتریس کواریانس $[n \times n]$ خوشه k ام است که بصورت زیر تعریف می‌شود:

$$\sum_k = \frac{1}{N_k} \sum_{t=1}^{N_k} I_k(x(t)) (x(t) - c_k) (x(t) - c_k)^T \quad (\text{ب-۹})$$

در این رابطه N_k تعداد کل نقاط متعلق به خوشه k ام است. اگر $x(t)$ به خوشه k ام تعلق داشته باشد آنگاه $I_k(x(t))$ برابر ۱ است در غیر اینصورت $I_k(x(t))$ برابر صفر خواهد بود.

البته استفاده از فاصله M در یادگیری رقابتی مشکلاتی نیز به همراه دارد بعنوان مثال رابطه شامل محاسبه معکوس ماتریس نمونه X می‌باشد که باعث بالا رفتن بار محاسباتی می‌شود. همچنین اگر تعداد نقاط متعلق به خوشه‌ای در مقایسه با بعد ورودی (n) کم باشند آنگاه ممکن است ماتریس کواریانس $[n \times n]$ بصورت منفرد درآید که باعث ایجاد مشکلاتی در محاسبه معکوس ماتریس خواهد شد.

^۱. Hyperellipsoidal

^۲. Mahalanobis Distance

برای حل مشکل منفرد شدن ماتریس خوشه‌ها، ابتدا خوشه‌ها را بصورت کروی در نظر می‌گیریم یعنی ماتریس کواریانس هر کدام از خوشه‌ها بصورت ماتریس واحد فرض می‌شود. همچنین جهت کاهش حجم محاسبات معکوس کردن ماتریسها، رابطه‌ای برای بروز رسانی مستقیم معکوس این ماتریسها بدست می‌آوریم. با توجه به تعریف نوروں برنده و نوروں رقیب ، با وارد شده هر داده ماتریس کواریانس نوروں برنده بصورت زیر اصلاح می‌شود :

$$\begin{aligned}\sum_w &= \sum_w + \alpha_w \left((X - c_w) (X - c_w)^T - \sum_w \right) \\ &= (1 - \alpha_w) \sum_w + \alpha_w \left((X - c_w) (X - c_w)^T \right)\end{aligned}\quad (10-ب)$$

برای نوروں رقیب نیز خواهیم داشت :

$$\begin{aligned}\sum_r &= \sum_r - \alpha_r \left((X - c_r) (X - c_r)^T - \sum_r \right) \\ &= (1 - \alpha_r) \sum_r + \alpha_r \left((X - c_r) (X - c_r)^T \right)\end{aligned}\quad (11-ب)$$

که α_w و α_r به ترتیب سرعت‌های یادگیری نوروںهای برنده و رقیب می‌باشند. سپس پارامترهای زیر را تعریف می‌کنیم :

$$\begin{aligned}m_k &= (x - c_k) \\ \lambda_k &= m_k^T \sum_r^{-1} m_k\end{aligned}\quad (12-ب)$$

برای بروز رسانی معکوس ماتریسهای کواریانس ، از رابطه زیر استفاده می‌شود :

$$\sum_{n=0}^{\infty} (-\beta x)^{-1} \approx (1 + \beta x)^{-1}\quad (13-ب)$$

با استفاده از روابط (ب-۱۰) و (ب-۱۳) برای نوروں برنده خواهیم داشت :

$$\begin{aligned}\sum_w^{-1} &= \left((1 - \alpha_w) \sum_w + \alpha_w (m_w \cdot m_w^T) \right)^{-1} \\ &= \frac{\sum_w^{-1}}{1 - \alpha_w} \left(I + \frac{\alpha_w}{1 - \alpha_w} \sum_w^{-1} (m_w \cdot m_w^T) \right)^{-1} \\ &= \frac{\sum_w^{-1}}{1 - \alpha_w} \left(I + \beta_w \sum_w^{-1} (m_w \cdot m_w^T) \right)^{-1}\end{aligned}$$

$$\begin{aligned}
\Sigma_w^{-1} &= \frac{\Sigma_w^{-1}}{1-\alpha_w} \left(I - \beta_w \Sigma_w^{-1} (m_w \cdot m_w^T) + \beta_w^2 \Sigma_w^{-1} (m_w \cdot m_w^T) - \beta_w^3 \dots \right) \\
&= \frac{\Sigma_w^{-1}}{1-\alpha_w} \left(I - \beta_w \Sigma_w^{-1} (m_w \cdot m_w^T) \left(1 - \beta_w \lambda_w + \beta_w^2 \lambda_w^2 - \dots \right) \right) \\
&= \frac{\Sigma_w^{-1}}{1-\alpha_w} \left(I - \beta_w \Sigma_w^{-1} (m_w \cdot m_w^T) \left(\frac{1}{1 + \beta_w \lambda_w} \right) \right) \quad (14-ب) \\
&= \frac{\Sigma_w^{-1}}{1-\alpha_w} \frac{\beta_w}{1-\alpha_w} \frac{(\Sigma_w^{-1} m_w)^T \cdot (\Sigma_w^{-1} m_w)}{1 + \beta_w \lambda_w}
\end{aligned}$$

با انجام عملیات مشابهی، رابطه زیر برای بروزرسانی ماتریس کواریانس نوروں رقیب بدست می‌آید.

$$\Sigma_r^{-1} = \frac{\Sigma_r^{-1}}{1+\alpha_r} \frac{\beta_r}{1+\alpha_r} \frac{(\Sigma_r^{-1} m_r)^T \cdot (\Sigma_r^{-1} m_r)}{1 + \beta_r \lambda_r} \quad (15-ب)$$

الگوریتم یادگیری روش خوشه‌سازی فوق‌بیضوی همانند روش کروی می‌باشد و تنها تفاوت ایندو الگوریتم در تعریف فاصله می‌باشد.

C- الگوریتم خوشه‌سازی گوستافسون-کسل

مراحل این الگوریتم بصورت زیر می‌باشند:

- ۱- انتخاب تعداد خوشه‌ها (R)، توان وزنی^۱ ($m > 1$) و تیرانس خطا (ϵ)، (m غالباً برابر ۲ فرض می‌شود). همچنین ماتریس بخش‌بندی μ مقداردهی اولیه می‌شود.
- ۲- محاسبه موقعیت مراکز خوشه‌ها (l شماره تکرار را نشان می‌دهد)

$$c_i^l = \frac{\sum (\mu_{ik}^{(l-1)})^m z_k}{\sum (\mu_{ik}^{(l-1)})^m}, \quad 1 < i < R \quad (16-ب)$$

- ۳- محاسبه ماتریس کواریانس فازی

^۱ . Weighting Exponent

$$F_i = \frac{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m (z_k - c_i^l) (z_k - c_i^l)^T}{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m}, \quad 1 < i < R \quad (17-ب)$$

۴- محاسبه فاصله بین هر داده (z_k) تا مراکز خوشه‌ها.

$$M_i = \det(F_i)^{1/n} F_i^{-1}, \quad 1 < i < R \quad (18-ب)$$

$$d_{ik}^2 = (z_k - c_i)^T M_i (z_k - c_i), \quad 1 < i < R, \quad 1 < k < N$$

۵- بروزرسانی ماتریس بخش‌بندی μ (محاسبه توابع عضویت خوشه‌ها)

$$\mu_{ik}^l = \frac{1}{\sum_{j=1}^R \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}}, \quad 1 < i < R, \quad 1 < k < N \quad (19-ب)$$

Abstract

This thesis is an applied research about the neuro-fuzzy generalized predictive controller and its application to a continuous stirred tank reactor (CSTR). Model predictive controllers have an extensive applications in industrial control due to their interesting characteristics. These controllers optimize the performance of a process and the practical constraints are considered directly in the solution of optimization problem. The concepts of model predictive methods are intuitive and they are applicable for a wide range of processes. These methods needs a predictive model to predict the future output of a process, which directly affects the performance of a predictive controller. Furthermore, an optimization problem should be solved in each sampling period. Identification of linear or static models is simple and the corresponding optimization problem has an analytical solution. Unfortunately, these models cannot capture the dynamic features of nonlinear time-variant systems and predict their future behaviour exactly. There are several nonlinear models which are capable to represent the dynamic systems, but these models are difficult to extract .More important problem is that the obtained optimization problem is nonlinear (and often non-convex). There is no analytical solution for such problem and it should be solved using complex and time-consuming methods. To solve these problems, two algorithms for online identification of a nonlinear fuzzy-adaptive Takagi-Sugeno model are presented after the description of basic concepts of a predictive controllers and a generalized predictive control scheme. The presented identification algorithms are applied to a benchmark modeling problems and it is shown that the obtained models have high accuracy, compact structure and an acceptable computational load. Using nonlinear Takagi-Sugeno fuzzy model, the difficulties of optimization problem reduced significantly because the sub-models of the TS model are linear. The nonlinear TS fuzzy model of a process under control is derived, updated, linearized instaneously and combined with a linear generalized predictive control scheme to obtain a fuzzy-adaptive generalized predictive controller. This controller is applied to a continuous stirred tank reactor and its performance in a different conditions (start -up, variable parameters and measurement noise) is analyzed.

KEYWORDS: Generalized Predictive Control – Clustering – Online Identification of Takagi-Sugeno Fuzzy Models – Online Optimization – Practical Constraints – Computational Load Reduction – Continuous Stirred Tank Reactor