



وزارت علوم، تحقیقات، و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی

طراحی الگوریتم‌های وفقی برای تفکیک مودی تعمیم یافته سیگنال‌های با قابلیت نمایش تنک

مجری

محمد مهدی جهانی یکتا

آذر ۱۳۹۲

بسم الله الرحمن الرحيم

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی «طراحی الگوریتم‌های وفقی برای تفکیک مودی تعمیم یافته سیگنالهای با قابلیت نمایش تنک» پیرو قرارداد شماره ۳۳/۱۴۶ مورخ ۱۳۹۲/۰۱/۱۷ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و آقای محمدمهدی جهانی یکتا (مجری) اجرا شده است.

این گزارش و تمام حقوق مادی آن بر اساس «قانون حمایت از حقوق مولفان و مصنفان و هنرمندان، مصوب سال ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران است و هرگونه استفاده از تمامی یا پاره‌ای از آن، شامل: نقل قول، تکثیر، انتشار، کاربرد نتایج، تکمیل، و مانند آنها به صورت چاپی، الکترونیکی، یا وسایل دیگر فقط با اجازه کتبی پژوهشگاه امکان‌پذیر است. نقل قول در حد هزار واژه در انتشارات علمی مانند کتاب و مقاله با درج اطلاعات کامل کتاب‌شناختی نیازی به مجوز پژوهشگاه ندارد. صحت مندرجات گزارش بر عهده مجری طرح پژوهشی است.

طراحی الگوریتم‌های وفقی برای تفکیک مودی تعمیم یافته سیگنال‌های با قابلیت نمایش تنک

مدیر طرح پژوهشی: محمدمهدی جهانی یکتا، پژوهشگاه علوم و فناوری اطلاعات ایران

نشانی: تهران، خیابان انقلاب، چهارراه فلسطین، شماره ۱۰۹۰

صندوق پستی: ۱۳۱۵۷۷۳۳۱۴

رایانامه: m_mahdi_jahani@yahoo.com

چکیده

در این پژوهش پس از مروری بر پیشینه و برخی الگوریتم‌های موجود برای تفکیک مودی تعمیم یافته، برخی از این الگوریتم‌ها را برای به کارگیری در تفکیک مودی سیگنال‌های تنک مورد تغییر و بهینه‌سازی قرار می‌دهیم. چند الگوریتم جدید نیز برای این کار ابداع نموده، و آنها را مورد تحلیل و بررسی قرار خواهیم داد. در ادامه، عملکرد الگوریتم‌های بهینه‌سازی یا ابداع شده در کاربردهایی واقعی روی سیگنال‌های تنک را مورد مطالعه قرار داده، و نشان خواهیم داد که خروجی آنها در عمل نیز بسیار مناسب است.

کلیدواژه‌ها: تفکیک مودی تعمیم یافته، سیگنال، مودهای اصلی

فهرست مطالب

۱	۱- تفکیک مودی عادی و تعمیم یافته
۲	۱-۱- مقدمه
۲	۱-۲- جایگاه تفکیک در پردازش اطلاعات
۴	۱-۳- سری‌های زمانی و مفاهیم مربوطه
۷	۱-۴- تفکیک بر اساس مودهای داده‌ها
۱۷	۱-۵- تصویر کردن مبتنی بر نمونه
۲۰	۱-۶- تفکیک مودی تعمیم یافته
۲۱	۱-۷- الگوریتم‌های تفکیک مودی تعمیم یافته
۲۲	۱-۸- معکوس‌سازی ماتریس کوواریانس
۲۵	۲- مسائل و سیستم‌های معکوس و متدهای رگولاسیون
۲۶	۲-۱- مقدمه
۲۶	۲-۲- سیستم‌های معکوس
۲۸	۲-۳- فضای دیتا - فضای سیگنال - نویز
۲۸	۲-۴- مسائل ناهمگون
۳۰	۲-۵- مسئله بازسازی سیگنال از روی نسخه فیلتر شده و نویزی آن
۳۲	۲-۶- روشی پیشنهادی برای رگولاسیون سیستم‌های توصیف شونده با سری ولترا
۴۰	۲-۷- رگولاتورهای دیگر
۴۰	۲-۸- خلاصه
۴۱	۳- روش‌های عددی در معکوس‌سازی سیستم‌ها
۴۲	۳-۱- مقدمه‌ای بر روش‌های مبتنی بر محاسبات عددی مورد استفاده در معکوس‌سازی سیستم
۴۲	۳-۲- روش قالب‌ها مناسب برای سیگنال‌های تنک
۴۷	۳-۳- تسریع‌کننده‌های چندجمله‌ای مناسب برای سیگنال‌های تنک
۴۸	۳-۴- تسریع به روش Chebyshev
۴۹	۳-۵- روش Conjugate Gradient
۵۹	۳-۶- تصویر کردن بر روی مجموعه‌های محدب مناسب برای سیگنال‌های تنک
۶۲	۳-۷- خلاصه
۶۳	۴- معکوس‌سازی با استفاده از تقریب‌های متوالی مناسب برای تفکیک مودی تعمیم یافته
۶۴	۴-۱- مقدمه
۶۴	۴-۲- چند شرط لازم و چند شرط کافی
۷۵	۴-۳- مزایای معکوس‌سازی سیستم‌ها با استفاده از روش تقریب‌های متوالی
۷۶	۴-۴- کاربردهای روش تقریب‌های متوالی
۷۹	۴-۵- جمع‌بندی
۸۰	۵- شبیه‌سازی‌ها

فهرست شکلها

۳	۱-۱ مراحل استخراج دانش
۶	۲-۱ میانگین درآمد سالانه هر خانواده در مناطق مختلف ایالات متحده
۶	۳-۱ میانگین دمای هوا در نیویورک بصورت ماهانه
۱۰	۴-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازپایی شده از تبدیل DFT
۱۱	۵-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازپایی شده از تبدیل DWT
۱۲	۶-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازپایی شده از تبدیل PAA
۱۳	۷-۱ انتخاب محورهای جدید برای داده های دو بعدی
۱۴	۸-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازپایی شده از تبدیل PCA
۱۵	۹-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازپایی شده از تبدیل RP
۱۷	۱۰-۱ یک سری زمانی به طول ۲
۱۷	۱۱-۱ نمایش یک سری زمانی به طول ۲ در یک فضای دو بعدی
۱۸	۱۲-۱ نگاشت یک سری زمانی از فضای دو بعدی به فضای یک بعدی
۲۰	۱۳-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازپایی شده از تبدیل SBP
۳۰	۱-۲ تصویر مات دریافتی پس از معکوس سازی (چپ) و تصویر بازسازی شده
۳۱	۲-۲ تصاویر بازسازی شده با استفاده از رگولاسیون تقریبی (چپ) و دقیق (راست)
۳۴	۳-۲ نمایش گرافیکی سیستم معکوس سیستم توصیف شونده با سری Volterra
۳۷	۴-۲ بلوک دیاگرام pre linearizer سیستم توصیف شونده با سری Volterra
۳۹	۵-۲ سیگنالهای خطای متوالی در بازسازی یک سیگنال از روی نسخه عبوری آن از سیستم ولترا
۴۶	۱-۳ سیگنال اولیه به همراه بازسازی شده آن بعد از ۳ تکرار (چپ) - بعد از ۵ تکرار (راست)
۶۰	۲-۳ نمایش گرافیکی خاصیت دوم تصویر کردن بر روی مجموعه های محدب.
۶۲	۳-۳ تصویر کردن های متوالی بر روی مجموعه های محدب
۶۶	۴-۱ نمایش روش تقریب های متوالی به صورت گرافیکی
۷۱	۴-۲ نتیجه عملیات تقریب های متوالی برای $g(x) = x^2$ و $\lambda = 0.222$
۷۱	۴-۳ نتیجه عملیات تقریب های متوالی برای $g(x) = x^2$ و $\lambda = 0.1$
۷۲	۴-۴ نتیجه عملیات تقریب های متوالی برای $g(x) = x^2$ و $\lambda = 1.01$
۷۶	۵-۴ طیف سیگنال متناظر با پالس حوزه زمان
۷۶	۶-۴ طیف سیگنالی که اثرات نامطلوب کانولوشن غیر ایده آل را خنثی می نماید
۸۱	۵-۱ استفاده از معکوس سازی وفقی ماتریس
۸۲	۵-۲ مبتنی بر استفاده از مدل سازی ولترا
۸۲	۵-۳ مبتنی بر فیت کردن در مدل تئوری قالبها
۸۲	۵-۴ مبتنی بر روش تسریع چیشیف
۸۳	۵-۵ conjugate gradient مبتنی بر روش
۸۳	۵-۶ مبتنی بر روش تقریب های متوالی با مقدار پارامتر برابر ۰/۲
۸۳	۵-۷ مبتنی بر روش تقریب های متوالی با مقدار پارامتر برابر ۰/۴
۸۵	۵-۸ مبتنی بر روش تقریب های متوالی با مقدار پارامتر برابر ۰/۷

فصل اول

تفکیک مودی عادی و تعمیم یافته

۱. مقدمه

هدف این فصل بیان شرح مطالعات انجام شده برای این طرح پژوهشی است. مطالعات انجام شده شامل مطالبی در مورد سریهای زمانی، و روشهای مرسوم برای تفکیک مودی عادی و تعمیم یافته آنها است. از آنجایی که روشهای مربوط به تفکیک تعمیم یافته مبتنی بر روشهای مرسوم برای تفکیک مودی عادی است، در این فصل به بیان روشهای مرسوم برای تفکیک مودی عادی می پردازیم.

۲. جایگاه تفکیک در پردازش اطلاعات

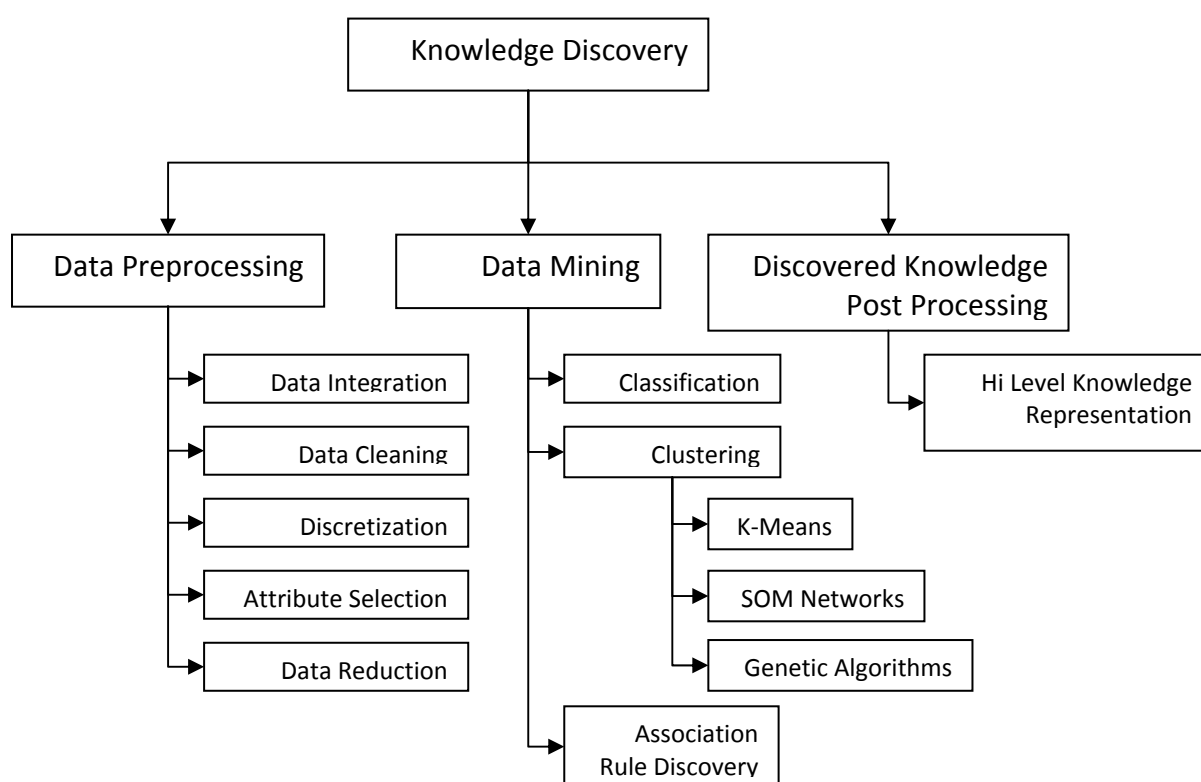
بسیاری از داده‌هایی که امروزه با آنها سروکار داریم داده‌هایی هستند که در طول زمان تغییر می‌کنند. دمای هوا، میزان مصرف انرژی در یک کشور، و ارتفاع سطح آب در یک رودخانه، همگی از نوع داده‌های وابسته به زمان هستند. بخش عمده‌ای از این داده‌ها از نوع سریهای زمانی می‌باشد. هرچند یک تعریف واحد برای سریهای زمانی وجود ندارد اما در بیشتر گزارشات تحقیقاتی و کتابها سری زمانی را ترتیبی از اعداد حقیقی در نظر می‌گیرند که مقدار مشاهده شده از یک رویداد را در فواصل زمانی برابر نشان می‌دهد.

دانشی که در این نوع داده‌ها نهفته است می‌تواند بسیار ارزشمند باشد. خوشبختانه امروزه با پیشرفتهایی که در علوم کامپیوتر بوجود آمده امکان استخراج این اطلاعات ارزشمند از داده‌ها پدید آمده است. استخراج این اطلاعات تحت فرایندی به نام استخراج دانش^۱ انجام می‌شود. استخراج دانش، طیف وسیعی از انواع داده‌ها را شامل می‌شود، مانند متن، جداول پایگاه داده، صفحات وب، سریهای زمانی و غیره. اینکه چه دانشی می‌خواهیم استخراج کنیم بستگی به نظر کاربر دارد. اما بطور کلی دانش استخراج شده باید صحیح، قابل درک و سودمند باشد. فرایند استخراج دانش را می‌توان به سه مرحله تقسیم نمود:

- پیش پردازش داده‌ها: در بیشتر مواقع قبل از اجرای الگوریتم‌های داده کاوی، لازم است که بر حسب نیاز، پیش‌پردازشهایی بر روی داده‌ها انجام شود. مانند یکی کردن منابع داده‌ها، پاکسازی داده‌ها، تبدیل مقادیر پیوسته به گسسته، انتخاب ویژگی‌های مناسب، کاهش حجم داده‌ها (مثلاً با استفاده از تکنیکهای کاهش ابعاد داده).

¹ Knowledge Discovery

- داده کاوی^۲: به معنی استخراج اتوماتیک (یا نیمه اتوماتیک) دانش از داده‌ها است. داده کاوی ترکیبی است از آمار، هوش مصنوعی، پایگاه داده و یادگیری ماشین. معروفترین تکنیکهای مورد استفاده در داده کاوی عبارتند از طبقه‌بندی^۳، خوشه‌بندی^۴ و استخراج قوانین انجمنی^۵.
 - پردازش دانش استخراج شده: مهمترین فعالیتی که در این مرحله انجام می‌شود نمایش دانش استخراج شده به کاربر است بطوریکه برای وی قابل درک باشد.
- این مراحل را در شکل ۱-۱ نشان داده‌ایم.



شکل ۱-۱ مراحل استخراج دانش

² Data Mining

³ Classification

⁴ Clustering

⁵ Association Rule Discovery

برای رسیدن به این هدف سعی می‌شود داده‌ها در گروه‌هایی دسته‌بندی شوند بطوریکه تفاوت داده‌های درون یک گروه مینیمم شده و تفاوت داده‌هایی که در گروه‌های مختلف وجود دارند ماکسیمم شود.

البته تاکنون انواع مختلفی از تفکیک گروهی تعریف شده است. بعضی‌ها تفکیک را بصورت hard در نظر گرفته‌اند و بعضی دیگر آن را فازی تعریف کرده‌اند.^۱

علاوه بر موارد فوق، نحوه تعریف فواصل بین داده‌ها، در گروه‌بندی ایجاد شده تأثیرگذار خواهد بود. بعضی از فرمول‌های اندازه‌گیری فاصله بین داده‌ها، عمومی هستند مانند فاصله اقلیدسی، در حالی که بعضی دیگر معمولاً برای داده‌های خاص به کار می‌روند مانند فاصله کسینوسی که برای داده‌های متنی استفاده می‌شود و یا فاصله DTW و فاصله correlation که برای سریهای زمانی تعریف شده‌اند.

۳. سریهای زمانی و مفاهیم مربوطه

در این بخش سریهای زمانی را معرفی کرده و مفاهیم مربوط به آن که در بخشهای آتی به آن نیاز خواهیم داشت را توضیح خواهیم داد. در بخش ۳-۱ به معرفی سریهای زمانی با ذکر چند مثال می-پردازیم. در بسیاری از تکنیکهایی که برای استخراج سیگنال از سریهای زمانی به کار می‌رود به معیاری برای سنجش میزان شباهت (یا تفاوت) دو سری زمانی نیاز است. مثلاً در خوشه‌بندی سریهای زمانی ما می‌خواهیم سریهایی که شبیه به هم هستند در یک خوشه قرار گیرند، بنابراین باید معیاری برای اندازه‌گیری شباهت دو سری زمانی داشته باشیم. در این مورد صحبت خواهیم کرد. علاوه بر این قبل از استفاده از تکنیکهای داده کاوی بر روی سریهای زمانی معمولاً به پیش‌پردازشهایی مانند نرمالسازی نیاز است تا داده کاوی بطور مؤثرتری انجام شود. سریهای زمانی معمولاً دارای دو ویژگی مهم هستند: اولاً ابعاد داده‌ها بسیار زیاد است و ثانیاً بین مقادیر یک سری زمانی همبستگی نسبتاً زیادی وجود دارد. به همین دلیل می‌توان برای کاهش حجم محاسبات، از تکنیکهای کاهش بعد استفاده نمود.

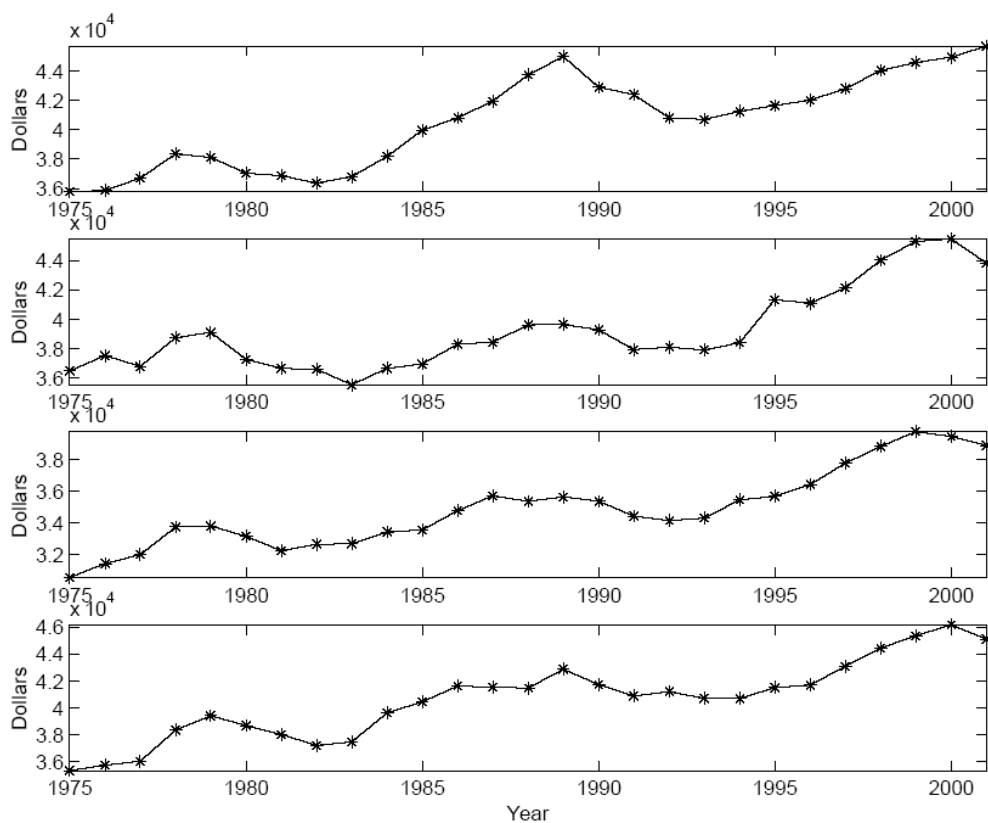
^۱ در خوشه‌بندی hard هر شیء فقط می‌تواند عضو یک خوشه باشد در حالی که در خوشه‌بندی فازی، هر شیء می‌تواند عضو چندین خوشه با درجه عضویت‌های مختلف باشد.

۳,۱. سریهای زمانی

داده‌هایی که از مشاهدات یک پدیده در طول زمان بدست می‌آیند بسیار متداول هستند. در هواشناسی بیشترین و کمترین درجه حرارت روزانه، باران سالانه و شاخصهای خشکسالی و سرعتهای باد در ساعات مختلف را مشاهده می‌کنیم. در علوم بیولوژیکی فعالیت الکتریکی قلب را در فواصل میلی ثانیه مشاهده می‌کنیم. در تمامی این موارد ما با داده‌های سری زمانی سروکار داریم. یک سری زمانی ترتیبی از مقادیر (معمولاً اعداد حقیقی) است که معمولاً با فواصل زمانی برابر مثلاً سالانه، ماهانه، هفتگی، روزانه و یا حتی فواصل زمانی بسیار کوچک مانند میلی ثانیه، ثبت شده است. در ادامه چند نمونه از سریهای زمانی را مشاهده می‌کنیم.

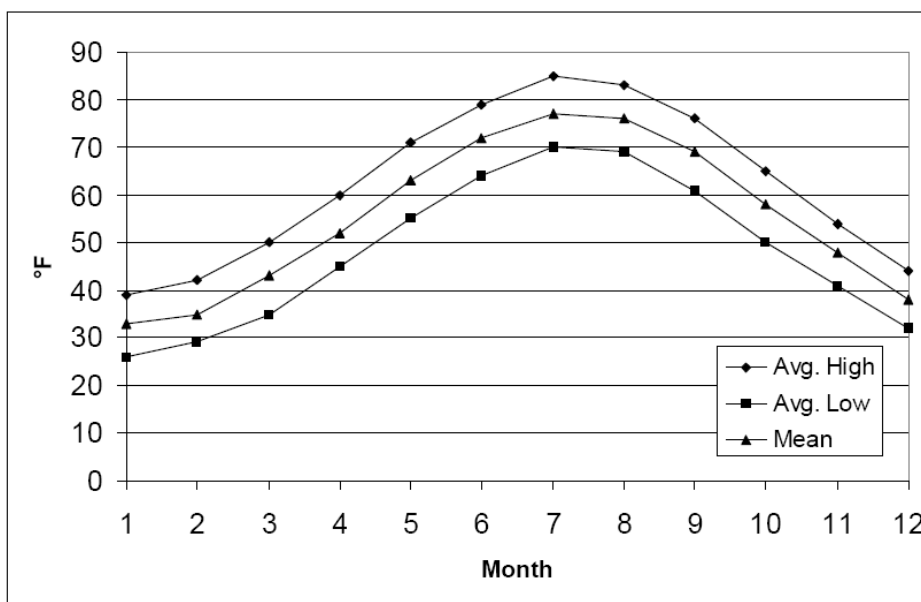
در شکل ۱-۲ میانگین درآمد سالانه هر خانواده در مناطق مختلف ایالات متحده از سال ۱۹۷۵ تا ۲۰۰۱ نشان داده شده است. اگر اقتصاد دانان بخواهند مسیر تغییرات درآمد سالانه را در طول زمان بررسی کنند، ارتباط بین سریهای زمانی مربوط به مناطق مختلف می‌تواند جالب باشد. در شکل ۱-۳ نیز میانگین دمای هوای نیویورک بصورت ماهانه نشان داده شده است.

سری زمانی S بصورت یک رشته مرتب از اعداد نشان داده می‌شود: $S = (x_1, x_2, \dots, x_n)$ که x_i مقدار مشاهده شده در زمان t_i می‌باشد. بدلیل اینکه این فاصله‌های زمانی معمولاً با هم برابر هستند، از نمایش آنها صرف نظر می‌شود و فقط مقادیر مشاهده شده را بصورت یک رشته از اعداد نشان می‌دهند.



شکل ۱-۲ میانگین درآمد سالانه هر خانواده در مناطق مختلف ایالات متحده از سال ۱۹۷۵ تا ۲۰۰۱، مناطق به

ترتیب از بالا به پایین: شمال شرقی، مرکز، جنوب، غرب



شکل ۱-۳ میانگین دمای هوا در نیویورک بصورت ماهانه

۳.۲. معیارهای اندازه گیری فاصله

معیارهای اندازه گیری فاصله به ما کمک می کنند تا میزان شباهت (یا اختلاف) بین سریهای زمانی را اندازه گیری نماییم. تا کنون معیارهای متعددی برای اندازه گیری فاصله ارائه شده است که هر یک کاربرد خاص خود را دارد. از مهمترین آنها می توان به فاصله اقلیدسی که کاربردهای بسیار وسیعی دارد، فاصله cosine measure که معمولاً در داده کاوی از داده های متنی به کار می رود، فاصله Autocorrelation که آن نیز کاربردهای وسیعی دارد، و فاصله dynamic time warping که معمولاً در کاربردهای تشخیص گفتار به کار می رود، اشاره کرد. از میان این معیارها، فاصله اقلیدسی بدلیل برخورداری از عملکرد قابل قبول، و فهم و پیاده سازی آسان، دارای محبوبیت بیشتری نسبت به بقیه روشها است. در ادامه به توضیح این معیار می پردازیم.

فاصله اقلیدسی حالت خاصی از فاصله Minkowski است. سری های زمانی $A = (a_1, a_2, \dots, a_n)$ و $B = (b_1, b_2, \dots, b_n)$ را در نظر بگیرید. فاصله Minkowski این دو سری از رابطه زیر محاسبه می شود.

$$D(A, B) = \sqrt[p]{\sum_{i=1}^n (a_i - b_i)^p}$$

به ازای بعضی از مقادیر p معیارهای معروفی بدست می آید که در زیر آنها را نشان داده ایم.

$$p = 1 \quad D(A, B) = \sum_{i=1}^n |a_i - b_i| \quad \text{Manhattan Distance}$$

$$p = 2 \quad D(A, B) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad \text{Euclidean Distance}$$

$$p = \infty \quad D(A, B) = \max \{a_i - b_i \mid i = 1, 2, \dots, n\} \quad \text{Supremum Distance}$$

۴. تفکیک بر اساس مودهای داده ها

از دیدگاه استخراج اطلاعات، سری های زمانی از دو ویژگی مهم برخوردارند:

۱. ابعاد بالا: اگر هر مؤلفه از سری زمانی را معادل با یک بُعد در نظر بگیریم، یک سری زمانی معادل است با یک نقطه در فضایی با ابعاد بسیار بالا. یعنی یک سری زمانی به طول ۱۰۰۰ معادل است با یک نقطه در فضای ۱۰۰۰ بعدی. در عمل، داشتن سری هایی با طول ۱۰۰۰ امری عادی است اما پردازش داده های ۱۰۰۰ بعدی حتی با کامپیوترهای پیشرفته امروزی می تواند بسیار مشکل باشد.

۲. همبستگی مقادیر: خوشبختانه مقادیر متوالی از یک سری زمانی به هم وابسته هستند زیرا این مقادیر بر اساس اطلاعاتی است که در طول زمان از دنیای خارج دریافت می‌شود. این وابستگی بین مقادیر نزدیک به هم در یک سری زمانی، باعث ایجاد افزونگی در داده‌ها می‌شود و امکان استفاده از تکنیکهای کاهش داده را فراهم می‌آورد.

کاهش داده یکی از مفاهیم مهم استخراج اطلاعات است. تکنیکهای کاهش داده، حجم بزرگی از داده‌ها را به داده‌هایی خلاصه و قابل مدیریت تبدیل می‌کنند ضمن اینکه خصوصیات داده‌ها را تا جایی که امکان دارد حفظ می‌کنند. با توجه به دو ویژگی ذکر شده در مورد سری‌های زمانی، کاهش داده می‌تواند در داده‌های سری زمانی نیز بسیار مفید باشد و ابعاد آن را به میزان قابل توجهی کاهش دهد. مسلماً پردازش سری‌هایی با ابعاد ۴ یا ۵ بسیار راحت‌تر از پردازش سری‌هایی با ابعاد ۱۰۰۰ می‌باشد.

بسیاری از تکنیکهایی که برای کاهش داده استفاده می‌شوند را می‌توان برای سری‌های زمانی نیز به کار برد. از مهمترین این تکنیکها می‌توان به موارد زیر اشاره کرد:

- Discrete Fourier Transform (DFT)
- Discrete Wavelet Transform (DWT)
- Piecewise Aggregate Approximation (PAA)
- Principle Component Analysis (PCA)
- Random Projection (RP)
- Piecewise Linear Approximation (PLA)
- Adaptive Piecewise Constant Approximation (APCA)

۴.۱. تبدیل Discrete Fourier Transform (DFT)

تبدیل فوری نوعی تبدیل است که یک تابع را بصورت توابع پایه‌ای سینوسی که هر کدام در مقادیری ضرب شده‌اند نشان می‌دهد. از تبدیل فوری در بسیاری از زمینه‌های علمی مانند فیزیک، هندسه، آمار و پردازش سیگنال استفاده می‌شود. تبدیل فوری یک تبدیل برگشت پذیر است. این تبدیل می‌تواند به دو صورت پیوسته یا گسسته انجام شود. در کامپیوتر و بخصوص در پردازش سیگنال معمولاً از تبدیل فوری گسسته (DFT) استفاده می‌شود [۱].

تعریف DFT - بردار $\vec{x} = x_0, x_1, \dots, x_{n-1}$ را در نظر بگیرید. تبدیل فوری گسسته این بردار از رابطه‌ی زیر بدست می‌آید.

$$\vec{X} = DFT(\vec{x}) = (X_0, X_1, \dots, X_{n-1})$$

که در آن:

$$X_F = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} x_i e^{-j2\pi Fi/n} \quad F = 0, 1, \dots, n-1 \quad j = \sqrt{-1}$$

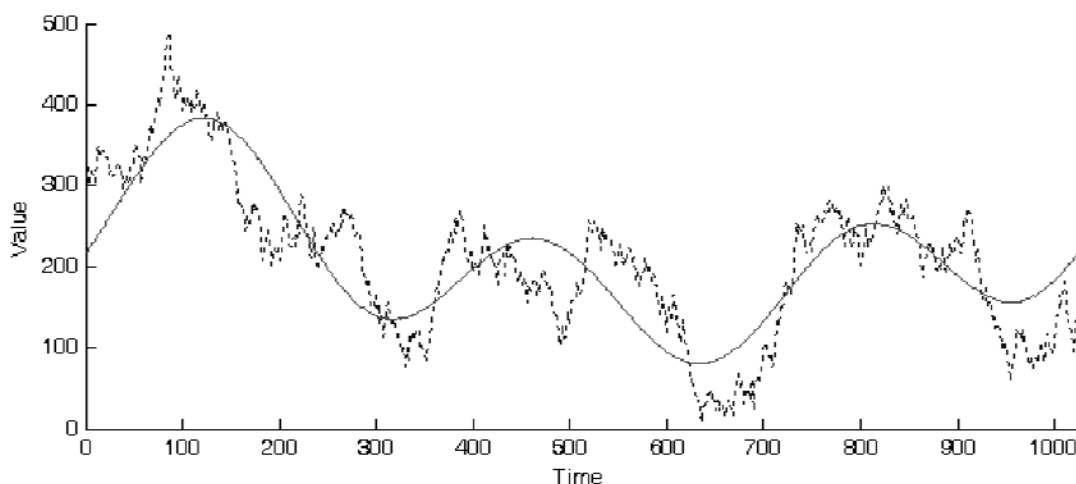
معکوس تبدیل فوریه گسسته نیز از این رابطه بدست می‌آید:

$$x_i = \frac{1}{\sqrt{n}} \sum_{F=0}^{n-1} X_F e^{j2\pi Fi/n} \quad i = 0, 1, \dots, n-1$$

توجه داشته باشید که دو بردار $\bar{X}$ و $\bar{x}$ هم‌اندازه هستند.

اگر بردار $\bar{x}$ را مقادیر یک سری زمانی در نظر بگیریم، که x_t مقدار مشاهده شده در زمان t است، تبدیل فوریه بر روی این بردار، مقادیر آن را از دامنه‌ی زمانی به دامنه‌ی فرکانسی منتقل می‌کند. ضرایب بدست آمده از تبدیل فوریه به گونه‌ای مرتب شده قرار می‌گیرند بطوریکه ضرایب پر اهمیت در سمت چپ و ضرایب کم اهمیت‌تر در سمت راست قرار می‌گیرند. بنابراین پر اهمیت‌ترین ضریب در X_0 و کم اهمیت‌ترین ضریب در X_{n-1} ذخیره شده است. هرچه یک ضریب پر اهمیت‌تر باشد، بدان معنی است که آن ضریب حاوی اطلاعات مهمتری (کلی‌تری) از بردار $\bar{x}$ می‌باشد. برای کاهش ابعاد داده $\bar{x}$ ، کافی است ابتدا $\bar{X} = DFT(\bar{x})$ را بدست آورده و سپس ضرایب کم اهمیت را حذف نمود.

در شکل زیر یک سری زمانی به طول ۱۰۲۴ را که با نقطه چین نشان داده شده مشاهده می‌کنید. ما با استفاده از تبدیل DFT ابعاد آن را به ۸ بعد کاهش دادیم و سپس سری اصلی را به کمک این ۸ مقدار بازیابی نمودیم. همانطور که مشاهده می‌شود سری بازیابی شده شبیه به یک نمودار سینوسی است.



شکل ۴-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده از تبدیل DFT که ابعاد آن را به ۸ بعد کاهش داده است

پیچیدگی زمانی محاسبه DFT برای یک سری زمانی به طول n برابر است با $O(n^2)$ ، اما خوشبختانه الگوریتم‌های سریعی تحت عنوان FFT^7 برای تبدیل فوری‌ی گسسته به وجود آمده است که دارای پیچیدگی زمانی $O(n \log n)$ می‌باشند.

۴.۲. Discrete Wavelet Transform (DWT)

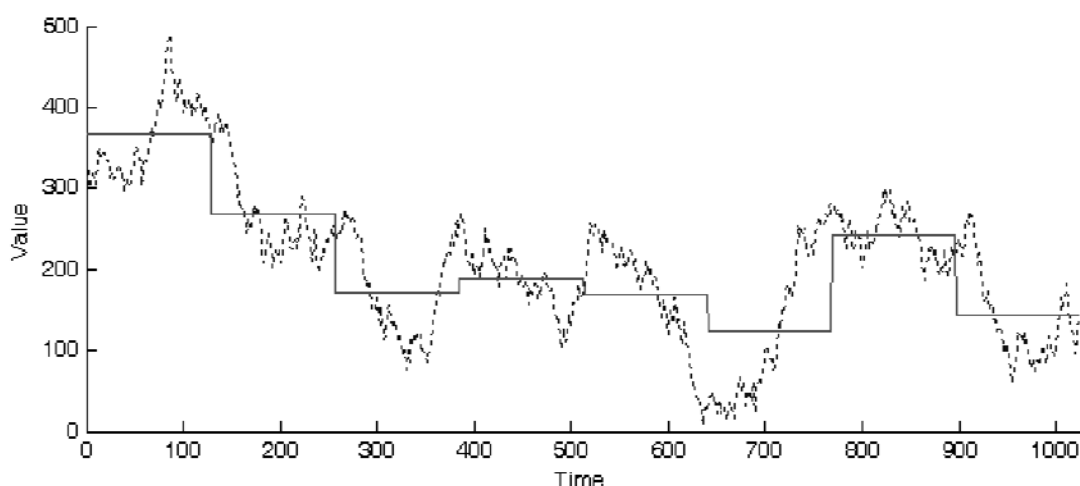
تبدیل DWT برای اولین بار توسط شخصی به نام Alfred Haar بوجود آمد. تا کنون نسخه‌های مختلفی برای DWT ارائه شده است، مانند Haar Wavelet، Newland Transform و Undecimated Wavelet Transform. این تبدیل نیز همانند تبدیل فوری‌ی بسیار پرکاربرد است و در بسیاری از زمینه‌های علوم و مهندسی مورد توجه قرار گرفته است. تبدیل Haar Wavelet بدلیل سادگی در پیاده سازی و سرعت اجرای بالا، از محبوبیت بیشتری نسبت به سایر نسخه‌های DWT برخوردار است [۲].

این تبدیل به اینصورت است که یک توالی به طول 2^n در ورودی داریم. این اعداد بصورت جفت جفت با هم جمع شده و این حاصل جمع‌ها به مرحله‌ی بعد فرستاده می‌شوند. همچنین اختلاف هر جفت نیز محاسبه و ذخیره می‌شود. دوباره این مرحله تکرار می‌شود با این تفاوت که در ورودی، حاصل جمع جفتهای مرحله‌ی قبل قرار می‌گیرد. این فرایند بطور بازگشتی تکرار می‌شود تا در نهایت یک عدد که حاصل جمع کل اعداد است بدست آید. این عدد به همراه 2^{n-1} اختلاف جفتهای که در

⁷ Fast Fourier Transform

مراحل مختلف الگوریتم محاسبه شده بعنوان خروجی این تبدیل بازگردانده می‌شود. پیچیدگی زمانی این الگوریتم برای یک رشته به طول n برابر با $O(n)$ می‌باشد.

اما چگونه می‌توان با استفاده از تبدیل DWT ابعاد داده را کاهش داد؟ در اینجا نیز همانند تبدیل فوریه، ضرایب بدست آمده به ترتیب بر اهمیت تا کم اهمیت مرتب شده‌اند. در واقع ضرایب کم اهمیت همانهایی هستند که در مراحل اولیه الگوریتم بدست می‌آیند. با حذف ضرایب کم اهمیت می‌توان حجم داده‌ها را کاهش داد. البته مقدار کمی از اطلاعات نیز از بین می‌رود. در شکل زیر یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده از ۸ ضریب پر اهمیت‌تر که از تبدیل DWT بدست آمده را مشاهده می‌کنید.



شکل ۵-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده از تبدیل DWT که ابعاد آن را به ۸ بعد کاهش داده است

۴.۳. Piecewise Aggregate Approximation (PAA)

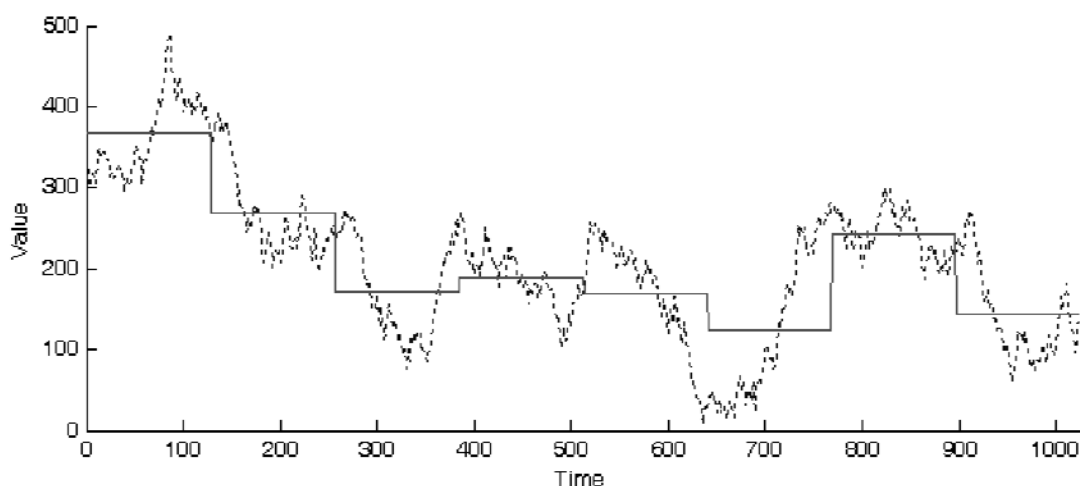
روش PAA برای اولین بار به منظور استفاده در حل مسأله جستجوی شباهت^۸ مطرح شد [۳]. در این روش برای کاهش ابعاد یک سری زمانی از n به N ، آن را به N توالی مساوی تقسیم کرده و میانگین هر توالی را بطور جداگانه محاسبه می‌کنند. اگر سری زمانی $\bar{x}$ را با $x_1, x_2, \dots, x_n$ نشان دهیم، سری زمانی کاهش یافته آن توسط PAA را با $\bar{X} = \bar{x}_1, \bar{x}_2, \dots, \bar{x}_N$ نشان می‌دهیم که:

⁸ Similarity Search

$$\bar{x}_i = \frac{N}{n} \sum_{j=\frac{n}{N}(i-1)+1}^{\frac{n}{N}i} x_j$$

واضح است که پیچیدگی زمانی این الگوریتم برابر است با $O(n)$. نشان داده شده که روش PAA علی‌رغم سادگی، کارایی آن در مسأله جستجوی شباهت به اندازه روشهای معروف دیگر کاهش ابعاد داده (نظیر DFT و DWT) می‌باشد.

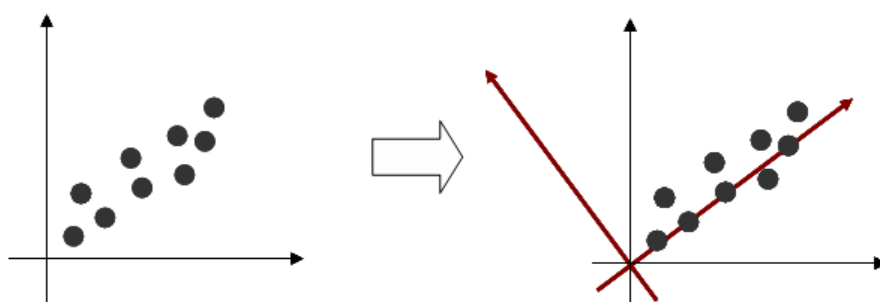
در شکل ۱-۶ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده که توسط PAA طول آن به ۸ بعد کاهش پیدا کرده بود نشان داده شده است. همانطور که مشاهده می‌کنید سری زمانی بازیابی شده که از روش PAA بدست می‌آید، بسیار شبیه به نتیجه بدست آمده از DWT است. اما باید توجه داشت که دو تفاوت عمده میان این دو روش وجود دارد. اول اینکه در ک الگوریتم و پیاده سازی PAA به مراتب ساده‌تر از DWT است و دوم اینکه در روش DWT طول سری زمانی باید توانی از ۲ باشد، در غیر اینصورت باید تعدادی صفر به انتهای مقادیر سری زمانی اضافه کرد تا طول آن توانی از ۲ شود. مسلماً این کار باعث هدر رفتن ظرفیت نگهداری اطلاعات و در نهایت کاهش کارایی این روش خواهد شد.



شکل ۱-۶ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده از تبدیل PAA که ابعاد آن را به ۸ بعد کاهش داده است

۴.۴. Principal Component Analysis (PCA)

تکنیک PCA بهترین روش برای کاهش ابعاد داده به صورت خطی می‌باشد [۴]. یعنی با حذف ضرایب کم‌اهمیت بدست آمده از این تبدیل، اطلاعات از دست رفته نسبت به روشهای دیگر کمتر است. البته کاربرد PCA محدود به کاهش ابعاد داده نمی‌شود و در زمینه‌های دیگری مانند شناسایی الگو و تشخیص چهره نیز مورد استفاده قرار می‌گیرد. در این روش محورهای مختصات جدیدی برای داده‌ها تعریف شده و داده‌ها براساس این محورهای مختصات جدید بیان می‌شوند. اولین محور باید در جهتی قرار گیرد که واریانس داده‌ها ماکسیمم شود (یعنی در جهتی که پراکندگی داده‌ها بیشتر است). دومین محور باید عمود بر محور اول به گونه‌ای قرار گیرد که واریانس داده‌ها ماکسیمم شود. به همین ترتیب محورهای بعدی عمود بر تمامی محورهای قبلی به گونه‌ای قرار می‌گیرند که داده‌ها در آن جهت دارای بیشترین پراکندگی باشند. در شکل زیر این مطلب برای داده‌های دو بعدی نشان داده شده است.

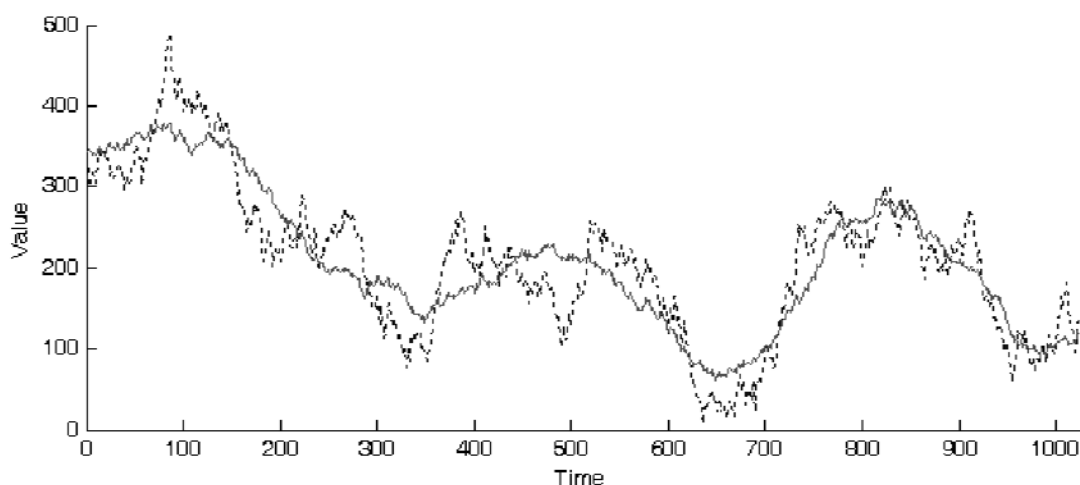


شکل ۱-۷ انتخاب محورهای جدید برای داده‌های دو بعدی

بر خلاف سه روش قبلی که کاهش ابعاد داده برای هر سری بطور جداگانه انجام می‌شد، در این روش ابتدا کل داده‌های موجود مورد بررسی قرار می‌گیرد تا وضعیت توزیع آنها مشخص شود. سپس محورهای مختصات جدید محاسبه شده و داده‌ها بر اساس این مختصات جدید بیان می‌شوند. یافتن محورهای مختصات جدید، از طریق محاسبه ماتریس کواریانس داده‌ها و سپس یافتن بردارهای ویژه-ی آن انجام می‌شود. توضیح کامل این روش خارج از حوصله‌ی این متن می‌باشد [۵].

در اینجا نیز بمنظور ارائه در کی بهتر نسبت به تأثیر این روش بر روی یک سری زمانی، در شکل زیر یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازایی شده که توسط PCA طول آن به ۸

بعد کاهش پیدا کرده بود را آورده‌ایم. همانطور که انتظار می‌رفت عملکرد این روش بهتر از سایر روش‌ها می‌باشد.



شکل ۸-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده از تبدیل PCA که ابعاد آن را به ۸ بعد کاهش داده است

مهمترین اشکال روش PCA مربوط به پیچیدگی زمانی آن است که برابر با $O(mn^2)$ می‌باشد. در اینجا m بیانگر تعداد داده‌ها (تعداد سری‌ها)، و n بیانگر ابعاد داده‌ها (طول سری‌ها) است. البته تاکنون روش‌هایی نیز برای افزایش سرعت محاسبات PCA ارائه شده‌اند که می‌توانند در زمان $O(mnN)$ ابعاد داده‌ها را به N بعد کاهش دهند که این کار را به کمک الگوریتم EM^9 انجام داده‌اند.

۴.۵. Random Projection (RP)

در روش RP داده‌هایی که در فضای n بعدی قرار دارند، به داده‌هایی در فضای N بعدی که $n > N$ ، نگاشت می‌شوند [۶]. این کار به کمک ماتریس $R_{n \times N}$ که مقادیر آن بصورت تصادفی ایجاد شده‌اند انجام می‌شود. البته مقادیر ماتریس تصادفی R باید طوری انتخاب شده باشند که هر ستون آن دارای طول واحد باشد. برای بدست آوردن داده‌های کاهش یافته^{۱۰} کافی است که ماتریس داده‌های اصلی را در ماتریس تصادفی R ضرب نماییم. اگر ماتریس داده‌های اصلی که از m سری

⁹ Expectation Maximization Algorithm

^{۱۰} در این گزارش به منظور اختصار، به جای عبارت "داده‌هایی با ابعاد کاهش یافته"، از عبارت "داده‌های کاهش یافته" استفاده شده است.

زمانی هر یک به طول n تشکیل شده را با $OriginalData_{m \times n}$ نمایش دهیم، ماتریس داده‌های کاهش یافته به طریق زیر محاسبه می‌شود.

$$ReducedData_{m \times N} = OriginalData_{m \times n} \bullet R_{n \times N}$$

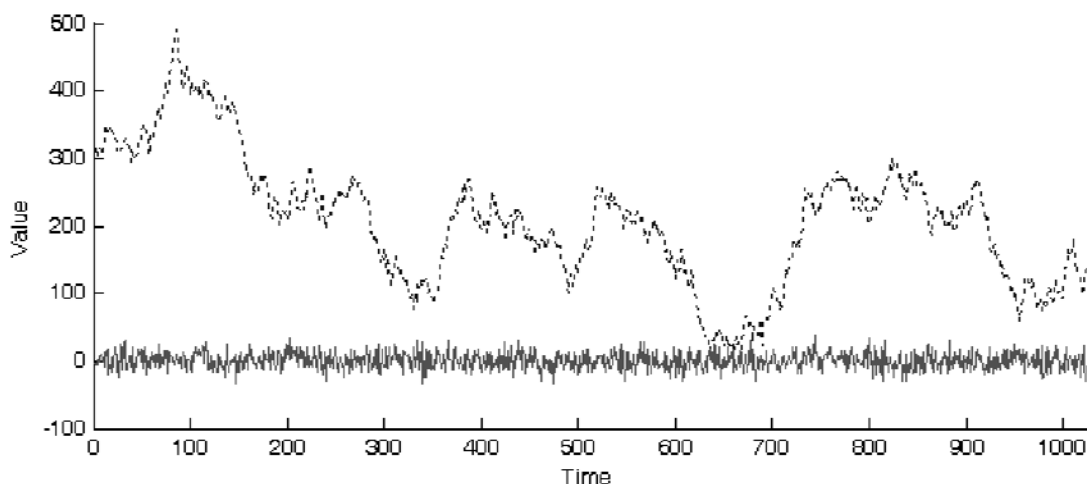
برای بازیابی داده‌های اصلی می‌توان از رابطه زیر استفاده کرد:

$$ReconstructedData_{m \times n} = ReducedData_{m \times N} \bullet R_{n \times N}^T$$

بنابراین پیچیدگی زمانی این روش $O(mnN)$ می‌باشد. مهمترین اشکال روش RP ناپایداری آن است، به این معنی که با هر بار اجرای RP بر روی یک مجموعه داده، نتایج مختلفی را دریافت می‌کنیم.

ایده روش RP نشأت گرفته از لم Johnson-Lindenstrauss می‌باشد که بیان می‌کند: به ازای هر $0 < \epsilon < 1$ ، هر مجموعه‌ای متشکل از m نقطه در فضای اقلیدسی n بعدی را می‌توان در فضایی با ابعاد $N = O(\log m / \epsilon^2)$ طوری جاسازی کرد که فواصل دو به دو بین تمام نقاط در فضای N بعدی، نسبت به فضای n بعدی، در بازه $(1 + \epsilon)$ قرار گیرد.

همانند چهار تکنیکی که در بالا شرح داده شد در اینجا نیز تأثیر این روش بر روی یک سری زمانی در شکل زیر نشان داده شده است. همانطور که مشاهده می‌کنید سری زمانی بازیابی شده، هیچ شباهتی با سری زمانی اصلی ندارد. اما از آنجایی که هدف از روشهای کاهش ابعاد داده معمولاً تفکیک است، و در تفکیک حفظ فواصل بین داده‌ها اهمیت دارد (و نه حفظ خود داده‌ها)، ضعف این روش در بازیابی سربهای زمانی اصلی، خللی در کار ایجاد نخواهد کرد.



شکل ۹-۱ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازایی شده از تبدیل RP که ابعاد آن را به ۸ بعد کاهش داده است

۴,۶. مقایسه تطبیقی روشهای کاهش ابعاد داده

در این بخش، پنج تکنیک کاهش ابعاد داده که در بالا مطرح شدند را با یکدیگر مقایسه می‌کنیم. این مقایسه در جدول زیر آورده شده است. پارامترهای مقایسه عبارتند از:

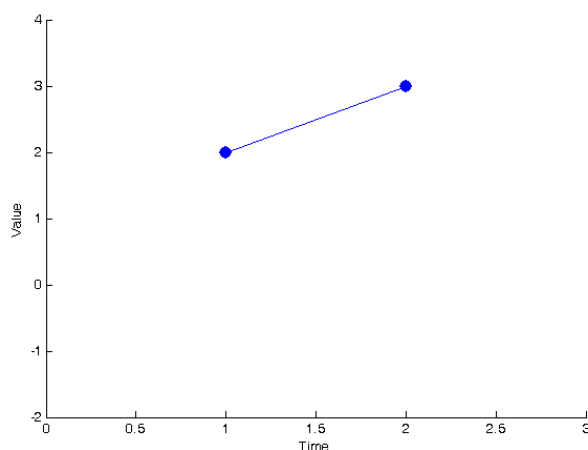
- نحوه تبدیل داده‌ها: همانطور که در بالا مشاهده شد، در بعضی از روشهای کاهش ابعاد داده، تبدیل هر داده (هر سری زمانی) بطور جداگانه و بدون توجه به نحوه توزیع داده‌ها انجام می‌شود اما در بعضی دیگر (مانند PCA) نحوه توزیع داده‌ها در نتیجه تبدیل، تأثیرگذار است. این دو روش را در ستون "نحوه تبدیل داده‌ها" به ترتیب با عبارات *independently* و *considering all data points* مشخص نموده‌ایم.
- پیچیدگی زمانی: در مورد پیچیدگی زمانی هر کدام از روشها در بالا صحبت شد و همچنین گفتیم که برای بعضی از این روشها الگوریتم‌های سریعی نیز ارائه شده است. پیچیدگی زمانی مربوط به الگوریتم معمولی و نیز الگوریتم سریع (در صورت وجود) در جدول نشان داده شده است. M بیانگر تعداد داده‌ها (سریها)، n بیانگر طول سریها، و N بیانگر طول سریها پس از کاهش ابعاد آنها می‌باشد.

مقایسه تطبیقی تکنیکهای کاهش ابعاد داده

تکنیکهای کاهش ابعاد داده	نحوه تبدیل داده‌ها	پیچیدگی زمانی	
		الگوریتم معمولی	الگوریتم سریع
DFT	Independently	$O(mn^2)$	$O(mn \log n)$
DWT	Independently	$O(mn)$	---
PAA	Independently	$O(mn)$	---
PCA	Considering all data points	$O(mn^2)$	$O(mnN)$
RP	Independently	$O(mnN)$	---

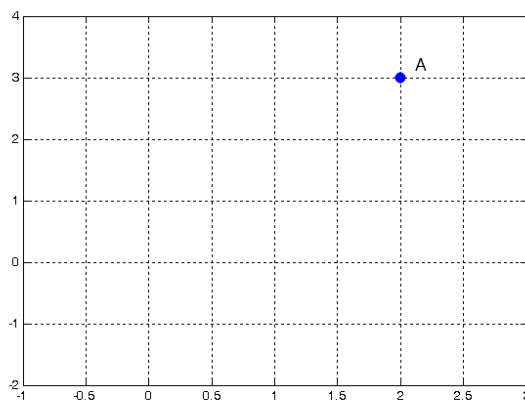
۵. Sample Based Projection (SBP)

همانطور که در بالا مشاهده شد سری زمانی بازیابی شده از داده‌های کاهش یافته توسط روش RP هیچ شباهتی با سری زمانی اصلی ندارد. علت این امر اینست که محورهای مختصات فضایی که قرار است داده‌ها در آن نگاشت شوند، بصورت تصادفی و بدون توجه به توزیع داده‌ها ایجاد شده‌اند. بعنوان مثال سری زمانی $A=(2,3)$ که یک سری زمانی به طول دو است را در نظر بگیرید (شکل ۱-۱۰). مقدار این سری در زمان $t=1$ برابر با ۲ و در زمان $t=2$ برابر با ۳ می‌باشد.



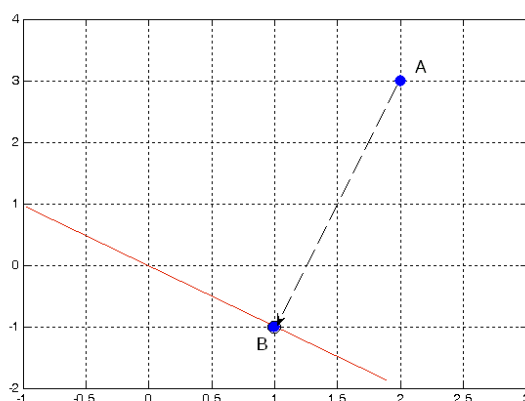
شکل ۱-۱۰ یک سری زمانی به طول ۲

می‌دانیم که یک سری زمانی به طول n را می‌توان بصورت یک نقطه در فضای n بعدی نمایش داد. در شکل ۱۱-۱ سری زمانی A را بصورت یک نقطه در فضای دو بعدی نشان داده‌ایم.



شکل ۱۱-۱ نمایش یک سری زمانی به طول ۲ در یک فضای دو بعدی

اکنون می‌خواهیم به وسیله روش RP طول این سری زمانی را از دو بُعد به یک بُعد کاهش دهیم. ابتدا باید یک فضای یک بُعدی که محور مختصات آن بصورت تصادفی ایجاد شده را بوجود آوریم. این فضا بصورت یک خط قرمز رنگ که از نقطه $(0,0)$ می‌گذرد در شکل ۱۲-۱ نشان داده شده است. سپس باید سری زمانی A را به این خط نگاشت کنیم. بدین ترتیب نقطه $B=(1,-1)$ را بدست خواهیم آورد. مختصات این نقطه در فضای جدید یک بُعدی (روی خط قرمز) برابر است با $\sqrt{5}$ که یک سری زمانی یک بُعدی است.



شکل ۱۲-۱ نگاشت یک سری زمانی از فضای دو بُعدی به فضای یک بُعدی

در روش RP محورهای مختصات بطور تصادفی ایجاد می‌شوند و پراکندگی داده‌ها هیچ نقشی در آنها ندارد. اما در روش SBP، این موضوع در ایجاد محورهای مختصات در نظر گرفته می‌شود [۷]. در این روش، ابتدا به ازای $i=1, \dots, n$ میانگین مؤلفه i ام تمام سریهای زمانی از مؤلفه i ام هر سری زمانی کم میشود (این کار در PCA نیز انجام می‌شود). بدین ترتیب میانگین تمام سریهای زمانی، رشته‌ای به طول n خواهد بود که تمام مقادیر آن صفر می‌باشد. برای کاهش ابعاد داده از n به N ، نیاز به N محور مختصات داریم که هر کدام توسط برداری با n مؤلفه و به طول واحد مشخص می‌شوند. برای بدست آوردن این N بردار، ابتدا N سری زمانی را بطور تصادفی از مجموعه کل سریها انتخاب می‌کنیم. سپس باید این N بردار را بر یکدیگر عمود سازیم. عمود سازی بردارها در روش RP چندان مورد نیاز نیست زیرا بردارهایی که بصورت تصادفی در یک فضای n بُعدی ایجاد می‌شوند، در صورت بزرگ بودن n تقریباً بر هم عمودند. اما در روش SBP بدلیل انتخاب بردارها از مجموعه داده‌ها، بردارها دیگر به طور تقریبی بر هم عمود نیستند. عمود سازی بردارها بر یکدیگر توسط روش

Gram-Schmidt قابل انجام است. بردارهای v_1 تا v_N را در نظر بگیرید که می‌خواهیم آنها را بر یکدیگر عمود سازیم. ابتدا عملگر $proj$ را بصورت زیر تعریف می‌کنیم.

$$proj_u^v = \frac{\langle v, u \rangle}{\langle u, u \rangle} u$$

که $\langle v, u \rangle$ ضرب داخلی بردارهای u و v را نشان می‌دهد. سپس بردارهای u_1 تا u_N که بر یکدیگر عمود هستند را بصورت زیر بدست می‌آوریم.

$$u_1 = v_1$$

$$u_2 = v_2 - proj_{u_1}^{v_2}$$

$$u_3 = v_3 - proj_{u_1}^{v_3} - proj_{u_2}^{v_3}$$

$$\vdots$$

$$u_N = v_N - \sum_{i=1}^{N-1} proj_{u_i}^{v_N}$$

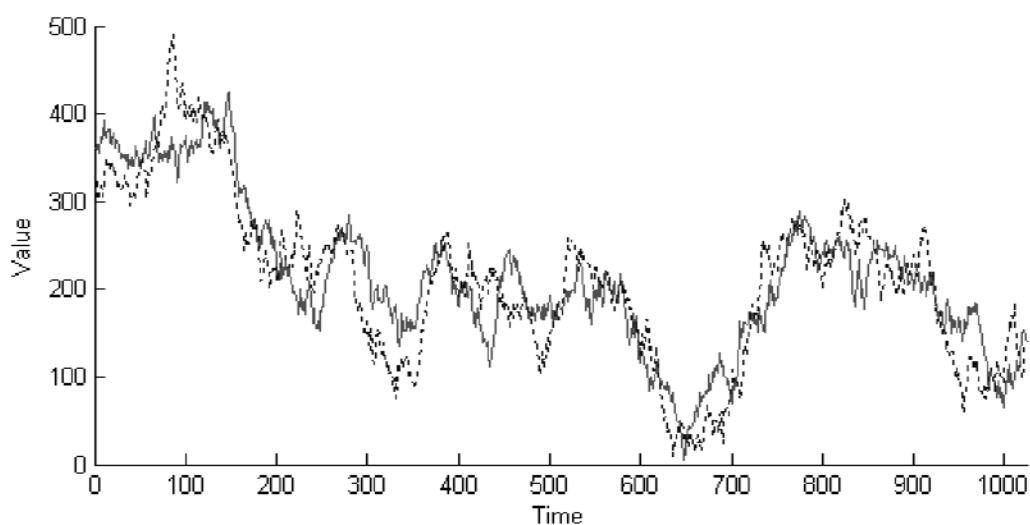
مرحله بعدی در روش SBP یکه کردن طول بردارهای عمود بر هم می‌باشد که با تقسیم هر بردار بر طول آن بدست می‌آید. بردارهای بدست آمده را در ستونهای ماتریس $U_{n \times N}$ ذخیره می‌کنیم. برای بدست آوردن داده‌های کاهش یافته کافی است که ماتریس داده‌های اصلی را در ماتریس U ضرب نماییم.

$$ReducedData_{m \times N} = OriginalData_{m \times n} \bullet U_{n \times N}$$

برای بازیابی داده‌های اصلی می‌توان از رابطه زیر استفاده کرد:

$$ReconstructedData_{m \times n} = ReducedData_{m \times N} \bullet U_{n \times N}^T$$

پیچیدگی زمانی این روش $O(nN^2 + mnN)$ می‌باشد. بنابراین برای N های کوچک، این روش بسیار مناسب است. در اینجا برای نشان دادن تأثیر این روش بر داده‌ها، یک سری زمانی را بوسیله SBP از ۱۰۲۴ به ۸ بعد کاهش داده و سپس آن را بازیابی نموده‌ایم. نتیجه را می‌توانید در شکل ۱-۱۳ مشاهده نمایید. در جدول ۲-۲ نیز مقایسه تطبیقی روش SBP را با تکنیکهای دیگر کاهش ابعاد داده مشاهده می‌کنید.



شکل ۱-۱۳ یک سری زمانی به طول ۱۰۲۴ به همراه سری زمانی بازیابی شده از تبدیل SBP که ابعاد آن را به ۸ بعد کاهش داده است

مقایسه تطبیقی روش SBP با تکنیکهای دیگر کاهش ابعاد داده

تکنیکهای کاهش ابعاد داده	نحوه تبدیل داده‌ها	پیچیدگی زمانی	
		الگوریتم معمولی	الگوریتم سریع
DFT	Independently	$O(mn^2)$	$O(mn \log n)$
DWT	Independently	$O(mn)$	---
PAA	Independently	$O(mn)$	---
PCA	Considering all data points	$O(mn^2)$	$O(mnN)$
RP	Independently	$O(mnN)$	---
SBP	Considering all data points	$O(nN^2 + mnN)$	---

۶. تفکیک مودی تعمیم یافته

انواع روشهای مورد استفاده برای تفکیک مودی در بخش‌های قبلی مورد بررسی قرار گرفتند. در روش تعمیم یافته، اصول کار همان اصول مندرج در بخشهای قبلی هستند؛ با این تفاوت که

ماتریسی که تفکیک بر اساس آن انجام می‌شود ماتریس همبستگی فرآیند اصلی نیست؛ بلکه ماتریس دیگری است که همبستگی فرآیندی غیر از فرآیند اول ولی وابسته به آن را توصیف می‌کند [۸].

به عنوان مثالی از سناریوهایی که باید از روش تعمیم یافته استفاده کرد، آنالیز روش کارکرد سیستم بینایی است. در این سیستم، تصاویر دریافت شده توسط مغز با استفاده از مودهای پردازشی قشر مغز به تعداد محدودی تصاویر پایه‌ای تبدیل شده، و ادراک بر حسب آنها صورت می‌گیرد [۹].

روشهای مهم برای تفکیک مودی تعمیم یافته عبارتند از روش گرادیان مزدوج، روش تصویر روی زیر فضاهای متعامد، روش فریم بندی، تفکیک لگاریتمی، و روش اپراتوری. این روشها در فصلهای بعدی به تفکیک مورد بررسی قرار می‌گیرند.

در بخش های قبلی الگوریتمهای متداول برای تفکیک مودی سیگنالها را مورد بررسی قرار دادیم. در ادامه، و طی فصول پیش رو به مطالعه الگوریتمهای موجود برای حالت تعمیم یافته می پردازیم و این الگوریتمها را برای استفاده در حالتی که سیگنالهای ورودی به آنها تنک هستند بهینه می‌نماییم. این بهینه سازی باعث می‌شود که از حجم حافظه و ظرفیت پردازشی پردازشگر به صورت بهتری استفاده گردد. در بسیاری از کاربردها که تمایل بر این است که محاسبات به صورت بلادرنگ در دسترس باشند چنین بهینه سازی هایی بسیار مورد نظر هستند.

ساختار ادامه گزارش به صورتی است که در پی می‌آید: در ادامه این فصل ارتباط تفکیک مودی عادی با تعمیم یافته مطرح می‌شود، و نشان داده می‌شود که چگونه با استفاده از یک سیستم معکوس، مسئله دوم به مسئله اول تبدیل می‌شود. در فصل ۲ به مسئله رگولارسازی می‌پردازیم که از طریق آن عملیات معکوس سازی به صورتی پایدار قابل انجام خواهد بود. مطالب این بخش به جز قسمتی که مربوط به معرفی سیستمهای ولترا است کار نویسنده است. در فصل ۳ به بررسی برخی از روشهای محاسباتی و عددی در معکوس سازی خواهیم پرداخت، که در آن، بخش معرفی تئوری قالبها و تسریع کننده ها برگرفته از منابع مرتبط بوده، و سایر قسمتها از ابتکارات نویسنده اند. بخشهای ۴ و ۵ نیز تماما حاصل کار نویسنده هستند.

۷. الگوریتمهای تفکیک مودی تعمیم یافته

در یک مسئله تفکیک مودی تعمیم یافته هدف آن است که یک سیگنال مثل X با ماتریس همبستگی A را بر حسب مودهای اصلی ماتریسی غیر از A که در اینجا آن را B می‌نامیم بیان نماییم. برای این کار، گام اصلی یافتن مقادیر ویژه و بردارهای ویژه تعمیم یافته است. و در صورتی که اینها

پیدا شوند بقیه کار همانند تفکیک مودی عادی است. لذا باید به دنبال مقادیر ویژه تعمیم یافته ای مثل S و بردار ویژه های تعمیم یافته ای مثل V باشیم به صورتی که داشته باشیم

$$AV=SBV \quad (1-1)$$

اگر B از ابتدا در دسترس بود به سادگی می توانستیم عبارت بالا را بصورت زیر بازنویسی کنیم

$$B^{-1}AV=SV \quad (2-1)$$

و مسئله تبدیل به یک مسئله تفکیک مودی عادی می شد. اما مسئله این است که سیگنالهای فرعی که ماتریس همبستگی B از روی آنها ساخته می شود از ابتدا در دسترس نیستند و در طی زمان به صورت یک به یک وارد سیستم می شوند. یعنی با ورود بردار n ام دیتا که آن را $u[n]$ می نامیم، تقریب n ام از B که آن را $B[n]$ می نامیم از روی تقریب قبلی آن به صورت زیر ساخته می شود:

$$B[n]=(1-a)*B[n-1]+a*u[n]u^T[n] \quad (3-1)$$

که در آن T برای نشان دادن ترانهاد ماتریس به کار رفته است، و a یک پارامتر کوچک مثلاً 0.01 است.

لذا مسئله اصلی تبدیل می شود به تخمین معکوس ماتریس B که خود در طول زمان تشکیل می شود. بنابراین یک مسئله تفکیک مودی تعمیم یافته را در واقع می توان به صورت یک مسئله معکوس سازی ماتریس تصور کرد که به دنبال آن یک مسئله تعمیم یافته عادی باید حل شود. لذا آن چیزی که باید در این گزارش در مورد آن بحث شود در واقع مسئله معکوس سازی سیستمی با ماتریس کوواریانس B است. بقیه مسئله هم که یک مسئله تفکیک مودی عادی است و در گزارش قبلی به صورت مبسوط مورد بررسی قرار گرفته است. لذا در ادامه به الگوریتمهای موجود برای معکوس سازی سیستمها می پردازیم. در گام اول، سیستم را به صورت خطی در نظر می گیریم، که معکوس سازی آن در واقع معادل با معکوس سازی ماتریس کوواریانس است.

۸. معکوس سازی ماتریس کوواریانس

فرض کنید $[A]$ و $[B]$ دو ماتریس معین مثبت $M \times M$ باشند، که این گونه به یکدیگر مرتبط

هستند :

$$[A] = [B]^{-1} + [C] [D]^{-1} [C]^H \quad (4-1)$$

که در آن، $[D]$ نیز یک ماتریس معین مثبت $M \times M$ و $[C]$ نیز یک ماتریس $M \times M$ می‌باشد. بنابراین لم معکوس کردن ماتریس‌ها، معکوس ماتریس $[A]$ این گونه بدست می‌آید:

$$[A]^{-1} = [B] - [B] [C] ([D] + [C]^H [B] [C])^{-1} [C]^H [B] \quad (5-1)$$

اگر برابری‌های زیر در معادلات قبل لحاظ شوند،

$$[A] = [\phi(n)] \quad (6-1)$$

$$[B^{-1}] = [\phi(n-1)]$$

$$[C] = \vec{u}(n)$$

$$[D] = 1$$

آنگاه طبق لم معکوس کردن ماتریس‌ها، نتایج زیر بدست می‌آیند:

$$[\phi^{-1}(n)] = [\phi^{-1}(n-1)] - \frac{[\phi^{-1}(n-1)] \vec{u}(n) \vec{u}^H(n) [\phi^{-1}(n-1)]}{1 + \vec{u}^H(n) [\phi^{-1}(n-1)] \vec{u}(n)} \quad (7-1)$$

برای سادگی محاسبات فرض کنید :

$$[p(n)] = [\phi^{-1}(n)] \quad (8-1)$$

و نیز :

$$\vec{k}(n) = \frac{[p(n-1)] \vec{u}(n)}{1 + \vec{u}^H(n) [p(n-1)] \vec{u}(n)} \quad (9-1)$$

با استفاده از این تعاریف، ماتریس معکوس قابل نوشتن به صورت زیر خواهد بود :

$$[p(n)] = [p(n-1)] - \vec{k}(n) \vec{u}^H(n) [p(n-1)] \quad (10-1)$$

به ماتریس $[p] (M \times M)$ ، ماتریس همبستگی معکوس گفته می‌شود. بردار $\vec{k} (M \times 1)$ را بردار بهره می‌نامیم. علت این نامگذاری بعداً روشن می‌شود. معادله اخیر را معادله ریکاتی در الگوریتم معکوس سازی می‌نامند.

از مرتب کردن عبارت مربوط به k داریم :

$$\begin{aligned}\vec{k}(n) &= [p(n-1)]\vec{u}(n) - \vec{k}(n)\vec{u}^H(n)[p(n-1)]\vec{u}(n) \\ &= [[p(n-1)] - \vec{k}(n)\vec{u}^H(n)[p(n-1)]]\vec{u}(n)\end{aligned}\quad (11-1)$$

می‌دانیم عبارت داخل براکت برابر با $[p(n)]$ است. پس داریم :

$$\vec{k}(n)[p(n)]\vec{u}(n) \quad (12-1)$$

این نتیجه، همراه با $[p(n)] = [\phi^{-1}(n)]$ می‌تواند به عنوان گام نهایی برای معکوس سازی استفاده شود.

در فصل بعد به بررسی مسائل و سیستمهای معکوس در حالت کلی تر آنها می‌پردازیم و متدهای رگولاسیون که در صورت تنک بودن سیگنالها قابل اعمال روی این روشها است را بررسی می‌کنیم.

فصل دوم

مسائل و سیستم‌های معکوس و متدهای رگولاسیون

۲-۱- مقدمه

در این فصل خواننده با مفهوم سیستم معکوس آشنا می‌گردد. به طور کلی هر روشی که برای بازیابی سیگنال ورودی یک سیستم از روی خروجی آن سیستم به کار رود سیستم معکوس یا معکوس ساز نامیده می‌شود. در بسیاری از موارد اگر سیستم معکوس به طور مستقیم بر روی خروجی سیستمی که هدف معکوس سازی آن است اعمال گردد، نمی‌تواند سیگنال اولیه را به طور قابل قبولی بازیابی نماید، اما اگر قبل از استفاده از آن، بر روی سیگنال خروجی این سیستم انواعی از پیش پردازش صورت گیرد و سپس از سیستم معکوس استفاده گردد، خروجی قابل قبول و با کیفیت بالا به دست می‌آید. در بسیاری دیگر از موارد، این نوع پیش پردازش‌ها به جای آن که بعد از سیستم صورت گیرند قبل از آن انجام می‌شوند. برای مثال در دنیای پردازش تصویر، با اطلاع از نوع اعوجاجی که در تصویر بر روی سیگنال ارسالی به وجود می‌آید، می‌توان پیش پردازش‌های لازم برای بهینه کردن عملکرد سیستم معکوس را انجام داد. در هر صورت، چه این که این پیش پردازش قبل یا بعد از سیستم انجام گیرد، به آن عادی سازی یا رگولاسیون^{۱۱} گفته می‌شود. یکی از مهمترین روش‌های رگولاسیون خطی سازی است. اگر یک معکوس ساز / خطی ساز قبل از سیستم اعوجاج قرار گیرد به آن پیش معکوس ساز / پیش خطی ساز^{۱۲} می‌گوییم و اگر بعد از آن قرار بگیرد به آن پس معکوس ساز / پس خطی ساز^{۱۳} گفته می‌شود.

در این فصل ابتدا مفهوم کلی سیستم معکوس بررسی می‌شود. سپس ایده رگولاسیون در قالب یک مثال مطرح شده، و بعد در مورد روش‌های رگولاسیونی که می‌توانند در مورد سیستم معکوس ساز به کار گرفته شود مطالبی ارائه می‌گردد.

۲-۲- سیستم‌های معکوس

هدف از جمع آوری دیتا به طور کلی به دست آوردن اطلاعات در مورد سیگنال‌های ارسالی و یا محتوای پدیده‌های فیزیکی است. اما در بسیاری از موارد سیگنال‌هایی که ما به دنبال آنها هستیم با سیگنال‌هایی که می‌توانیم آنها را اندازه گیری کنیم تفاوت دارند. استفاده از دیتاهای اندازه گیری شده برای به دست آوردن سیگنال‌های فرستاده شده به طور کلی یک مسئله معکوس، و سیستمی که

¹¹ Regulation

¹² pre equalizer / pre linearizer

¹³ post equalizer / post linearizer

این کار را انجام می‌دهد یک سیستم معکوس نامیده می‌شود. در این جا به معرفی چند نمونه از مسائل معکوس می‌پردازیم.

۲-۲-۱-توموگرافی محوری کامپیوتری^{۱۴}: در این نوع توموگرافی هدف به دست آوردن تصویر از یک مقطع درونی بدن است. اگر اشعه X ارسالی به درون بدن می‌توانست تنها از سطح مقطع مورد نظر تصویر برداری نماید، مسئله ای برای بررسی باقی نمی ماند، اما مشکلی که وجود این است که هر پرتو اثری بر روی فیلم رادیولوژی ایجاد می‌کند که حاصل انتگرال خطی از میزان عبور دهی بافت‌هایی است که پرتو از آنها گذشته است. مسئله معکوسی که مطرح است، بازسازی تصویر سطح مقطع مورد نظر از روی تصویر به دست آمده است.

۲-۲-۲- فیت کردن مدل: فرض کنید بنا بر یک مدل تئوریک داده‌های x و y از طریق رابطه زیر به هم مرتبط باشند:

$$y = a + bx + cx^2 + dx^3$$

هدف این است با چند اندازه گیری، پارامترهای مدل را تعیین نماییم. البته مدل مذکور می‌تواند بسیار پیچیده تر از چیزی باشد که در این مثال مطرح شد.

۲-۲-۳- Deconvolution: فرض کنید که یک سیگنال تصویر از یک محیط عبور کرده و دچار اعوجاج شده باشد، مثلاً مات و یا فیلتر شده باشد. یافتن سیستم‌هایی که بتوانند اثرات این اعوجاج را جبران نموده و تصویر شفاف را به دست دهند در حیطه مسائل معکوس می‌گنجد.

۲-۲-۴- تحلیل‌های عددی: به عنوان یک مسئله با صورت ریاضیاتی واضح تر، معادله انتگرالی زیر را در نظر بگیرید:

$$\int_a^b k(x,s)z(s)ds = y(x) \quad a \leq x \leq b$$

به این تساوی، معادله نوع اول Fredholm گفته می‌شود. فرض می‌شود کرنل $k(x,s)$ و تابع y معلوم اند و به دنبال یافتن تابع z می‌باشیم. این مسئله نیز جزو مسائل معکوس است.

از مطالب بیان شده واضح است که دامنه مسائل معکوس بسیار گسترده است، و حل آنها کاربردهای بسیار زیادی دارد. همان گونه که در مقدمه ذکر شد در برخی موارد نمی‌توان مستقیماً فر

¹⁴ Computer Axial Tomography

آیند معکوس سازی را آغاز نمود زیرا ممکن است سیستم معکوس به دست آمده ناپایدار، نوسانی، و یا تباهیده^{۱۵} شود. در این موارد قبل از شروع معکوس سازی لازم است روی سیگنال اعوجاج یافته پیش پردازش‌هایی به نام رگولاسیون صورت گیرد تا آماده ورود به پروسه معکوس سازی گردد. این موضوع را به صورت یک مثال بیشتر روشن می‌کنیم. این مثال مربوط است به مسئله باز یابی تصویر از روی نسخه مات شده آن. برای ارائه این مثال لازم است که خواننده ابتدا با چند مفهوم آشنا گردد.

۲-۳- فضای دیتا، فضای سیگنال، نویز

در مطالبی که در پی آورده می‌شود به مجموعه دیتاهای نادقیق در دست، " فضای دیتا"، به مجموعه همه سیگنال‌هایی که ممکن است در دست داشته باشیم فضای سیگنال، و به تفاوت دیتای در دست و سیگنال متناظر با آن نویز گفته می‌شود. سیگنال را با f نمایش می‌دهیم.

اگر سیستمی که باید معکوس شود را با A ، دیتا را با d ، و نویز را با n نمایش دهیم، رابطه زیر را بین این پارامترها داریم:

$$d = A(f) + n$$

با این تعاریف، مسئله معکوس عبارت است از یافتن سیگنال از روی دیتا با دانستن ویژگی‌های سیستمی که باید معکوس گردد.

۲-۴- مسائل ill-posed و ill-conditioned

طبق تعریف Hadamard در سال ۱۹۲۳، مسئله یافتن سیگنال از روی دیتا well-posed گفته می‌شود اگر سه شرط زیر برقرار باشند:

۱- برای هر دیتا یک سیگنال متناظر در فضای سیگنال وجود داشته باشد.

۲- سیگنال متناظر با هر دیتا، یکتا باشد.

۳- تناظر بین فضای دیتا به فضای سیگنال یک تناظر پیوسته باشد.

شروط ۱ و ۲ معادل این هستند که سیستم A یک معکوس خوش تعریف داشته باشد، که حوزه تعریف آن همه فضای دیتا باشد. شرط سوم یک شرط لازم و نه کافی برای پایداری سیستم معکوس است. برای مسائل well-posed، ارتباط بین میزان‌های نسبی خطا در فضای دیتا و فضای سیگنال از

¹⁵ Degenerate

طریق عدد حالت^{۱۶} به یکدیگر مربوط می‌شوند. اگر Δd و Δf به ترتیب برابر با میزان انحراف‌های دیتا و سیگنال از مقادیر واقعی خود باشند، عدد حالت χ چنان است که همواره داریم:

$$\frac{\|\Delta f\|}{\|f\|} \leq \chi \frac{\|\Delta d\|}{\|d\|}$$

برای سیستم‌ها خطی عدد حالت به صورت زیر تعیین می‌گردد:

$$\chi = \|A\| \|A^{-1}\|$$

در نظریه جبر خطی ثابت می‌گردد که [10]:

$$\chi = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

به بیان دیگر χ بیان گر حداکثر بزرگ شدن طول یک بردار یک‌تایی تحت سیستم خطی مربوطه است. واضح است که هر قدر عدد حالت کوچکتر باشد وضعیت سیستم معکوس از نظر میزان پایداری بهتر است. در مورد سیستم‌های خطی این عدد همواره بزرگتر یا مساوی ۱ است. در حالت کلی اگر عدد حالت یک سیستم کوچک باشد، مسئله یافتن سیستم معکوس آن را خوش حالت^{۱۷}، و در غیر این صورت آن را بد حالت^{۱۸} می‌نامند. روشن است که مرز خیلی دقیقی بین این دو مفهوم وجود ندارد.

Hadamard بر این باور بود که مسائل ill-posed اصولاً غیر واقعی بوده، و در دنیای فیزیکی رخ نمی‌دهند. این تصور اشتباه است. یک مثال نسبتاً ساده برای نشان دادن اشتباه بودن این تصور معادله Fredholm مرتبه اول با کرنل عضو فضای هیلبرت است:

$$\int_a^b k(x, s) f(s) ds = d(x) \quad a \leq x \leq b$$

اگر جواب این معادله با تابعی مانند تابع $\Delta f(s) = \varepsilon \sin(2\pi ps)$ ($p=1,2,\dots$) دچار اعوجاج گردد، آنگاه اعوجاجی که در دیتا رخ می‌دهد برابر خواهد بود با:

$$\Delta d(x) = \varepsilon \int_a^b k(x, s) \sin(2\pi ps) ds$$

¹⁶ Condition Number

¹⁷ Well-Conditioned

¹⁸ Ill-Conditioned

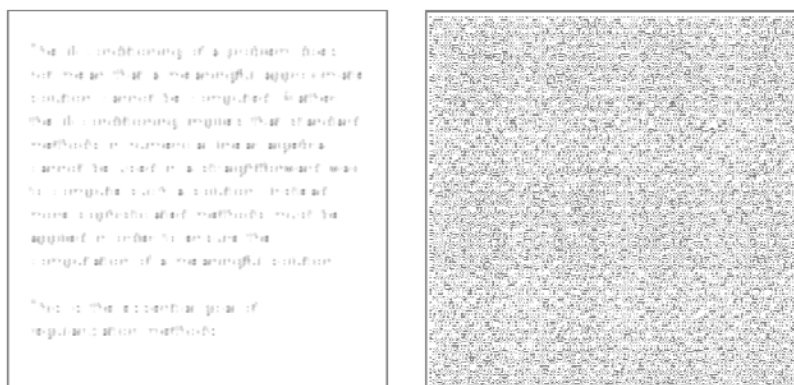
و بنابر قضیه ریمان در انتگرال‌ها روشن است که با شرط $p \rightarrow \infty$ داریم $\Delta d \rightarrow 0$. واضح است که می‌توان با انتخاب p به اندازه کافی بزرگ، مقدار $\frac{\|\Delta f\|}{\|\Delta d\|}$ را از هر عدد از قبل تعیین شده ای بزرگتر کرد و بنا بر این حل معادله انتگرالی فوق یک مسئله ill-posed است. این مسئله از آن جهت عنوان گردید که این تصور اشتباه که مسائل دنیای فیزیکی همیشه well-posed هستند برای خواننده به وجود نیاید.

۲-۵- مسئله بازسازی سیگنال از روی نسخه فیلتر شده و نویزی آن

اگر فرض کنیم که سیستمی که باید معکوس گردد تلفیق فیلتر کردن پایین گذر به همراه اضافه شدن نویز باشد، در حوزه فوریه می‌توان تبدیل انجام شده بر روی سیگنال را به صورت

$$D(\omega) = A(\omega)F(\omega) + N(\omega)$$

مدل سازی نمود. اگر نویز وجود نداشت، امکان بازیابی سیگنال با استفاده از رابطه $F(\omega) = \frac{D(\omega)}{A(\omega)}$ وجود داشت. اما اگر در عمل از این روش استفاده گردد، نتایج به دست آمده اصلاً قابل قبول نیستند. علت این پدیده حضور نویز پهن باند است. در فرکانس‌هایی که اندازه $A(\omega)$ بسیار کوچک است، نویز به شدت تقویت شده و کیفیت سیگنال باز سازی شده را به شدت پایین خواهد آورد. برای مثال یک سیگنال تصویر را در نظر بگیرید که به دلیل عبور از فیلتر پایین گذر مات شده است. اگر با استفاده از روش فوق سعی در بازسازی این تصویر نماییم، تصویر زیر به دست می‌آید:



شکل ۲-۱- تصویر مات دریافتی پس از معکوس سازی (چپ) و تصویر بازسازی شده از طریق فیلتر کردن بدون انجام رگولاسیون (راست)

با توجه به تصاویر فوق واضح است که باید از نوعی رگولاسیون برای بازسازی تصویر فوق استفاده کرد. یک روش برای انجام این کار استفاده از یک فیلتر پایین گذر است به نحوی که تمام مولفه‌های سیگنال را از خود عبور داده ولی مولفه‌های فرکانس بالای نویز را حذف کند. اگر قبل از اعمال سیستم معکوس از این روش رگولاسیون استفاده گردد، به شرط آن که فیلتر پایین گذری که به کار گرفته می‌شود به فیلتر ایده آل خیلی شبیه نباشد، سیگنال باز سازی شده نسبتاً شفاف، و اگر فیلتر مورد استفاده از درجه بالا تری باشد، سیگنال به دست آمده کاملاً دقیق و شفاف خواهد بود.

The ill-conditioning of a problem does not mean that a meaningful approximate solution cannot be computed. Rather the ill-conditioning implies that standard methods in numerical linear algebra cannot be used in a straightforward way to compute such a solution. Instead, more sophisticated methods must be applied in order to ensure the computation of a meaningful solution.

This is the essential goal of regularization methods.

The ill-conditioning of a problem does not mean that a meaningful approximate solution cannot be computed. Rather the ill-conditioning implies that standard methods in numerical linear algebra cannot be used in a straightforward way to compute such a solution. Instead, more sophisticated methods must be applied in order to ensure the computation of a meaningful solution.

This is the essential goal of regularization methods.

شکل ۲-۲- تصاویر بازسازی شده با استفاده از رگولاسیون تقریبی (چپ) و دقیق (راست)

تاکنون خواننده با مفهوم رگولاسیون و جایگاه آن در سیستم‌های معکوس آشنا شده است. در مورد روش‌های معکوس سازی و بالاخص روش معکوس سازی مورد استفاده در این پروژه نیز استفاده از رگولاسیون جایگاه ویژه ای دارد. مهمترین نوع رگولاسیونی که در انجام شبیه سازی‌ها از آن استفاده شده است، فیلتر کردن پایین گذر است. عدم استفاده از این نوع رگولاسیون در مورد برخی از سیستم‌ها می‌تواند به طور کلی باعث از کار افتادن روش گردد.

در فصول بعدی نشان می‌دهیم که روش‌های معکوس سازی مورد بحث می‌تواند تحت شرایط نسبتاً ساده ای وارون سیستم‌های خطی را بسازند، لذا انتظار می‌رود که معکوس سیستم‌هایی که رفتارشان به حالت خطی بسیار نزدیک است را نیز بتواند با تقریب‌های قابل قبولی ایجاد کنند. با توجه به این مطلب، واضح است که اگر بتوان معکوس یک سیستم را با انجام انواعی از پیش پردازش به صورت تقریبی ایجاد کرد، می‌توان کل سیستم (شامل سیستم اولیه و معکوس ساز تقریبی پس از آن) را به صورت یک سیستم تقریباً همانی در نظر گرفت که مسلماً تقریباً خطی نیز هست، بنابراین انتظار

می‌رود که تکنیک معکوس سازی بتواند بهبود قابل ملاحظه ای بر روی کیفیت خروجی آن ایجاد کند. این موضوع منجر به این می‌شود که بتوان هر گونه سیستمی که تقریبی از سیستم معکوس یک سیستم اولیه یا حداقل خطی ساز آن باشد را به عنوان یک رگولاتور برای روش معکوس سازی قلمداد کرد. از این رو مناسب است که روش‌های معکوس سازی و خطی سازی سیستم‌های متداول را مورد مطالعه قرار دهیم. سیستم B ای که بسیاری از انواع کوواریانسهای متداول در سیستم‌های پردازش تصویر و تفکیک مودی را مدل سازی می‌کند سیستم توصیف شونده با سری ولتر است. در این جا این نوع مدل سازی و خطی سازی آن را مورد بررسی قرار می‌دهیم.

۶-۲- روشی پیشنهادی برای رگولاسیون سیستم‌های توصیف شونده با سری‌های ولترا

یکی از متداول ترین سیستم‌های حاضر در بحث‌های تفکیک مودی، سیستم‌های قابل توصیف با سری‌های ولترا هستند [11] که خروجی آنها به صورت زیر به ورودی شان مرتبط می‌گردد:

$$y(n) = \sum_{i=0}^N a_i x(n-i) + \sum_{k=2}^L \sum_{i_1=1}^N \sum_{i_2=i_1}^N \dots \sum_{i_k=i_{k-1}}^N h_{i_1 i_2 \dots i_k} x(n-i_1) x(n-i_2) \dots x(n-i_k) \quad (1-2)$$

بحثی که ارائه خواهد شد قابل تعمیم به خانواده بسیار بزرگتری از سیستم‌های غیر خطی (شامل چند جمله ای‌های تقریب‌های متوالی) نیز هست. اما برای سادگی بحث در ادامه تنها سیستم‌های به فرم (۲-۱) را بررسی می‌کنیم.

۶-۲-۱- معکوس سازی سیستم‌های توصیف شونده با سری‌های ولترا

اگر سیگنال ورودی سیستم معکوس را با u و خروجی آن را با w نشان دهیم معکوس سیستم فوق توسط (۲-۲) به دست می‌آید.

$$w(n) = \frac{1}{a_0} [u(n) - \sum_{i=1}^N a_i w(n-i) + \sum_{k=2}^L \sum_{i_1=1}^N \sum_{i_2=i_1}^N \dots \sum_{i_k=i_{k-1}}^N h_{i_1 i_2 \dots i_k} w(n-i_1) w(n-i_2) \dots w(n-i_k)] \quad (2)$$

این که می‌توانیم معکوس سیستم غیر خطی مذکور را به صورت ساده ای مثل (۲-۲) نمایش دهیم به دلیل این است که خروجی سیستم غیر خطی در لحظه n به $x(n)$ بستگی غیر خطی ندارد. اگر چنین بستگی وجود داشته باشد $x(n)$ تنها می‌تواند به صورت ضمنی به دست آورده شود. در واقع اگر $y(n)$ به طور غیر خطی به $x(n)$ بستگی داشته باشد داریم:

$$y(n) = \sum_{i=0}^N a_i x(n-i) + \sum_{k=2}^L \sum_{i_1=0}^N \sum_{i_2=i_1}^N \dots \sum_{i_k=i_{k-1}}^N h_{i_1 i_2 \dots i_k} x(n-i_1) x(n-i_2) \dots x(n-i_k) \quad (2-)$$

(۳)

و در این صورت معکوس سیستم این چنین به دست می‌آید:

$$w(n) = \frac{1}{a_0} [u(n) - \sum_{i=1}^N a_i w(n-i) + \sum_{k=2}^L \sum_{i_1=0}^N \sum_{i_2=i_1}^N \dots \sum_{i_k=i_{k-1}}^N h_{i_1 i_2 \dots i_k} w(n-i_1) w(n-i_2) \dots w(n-i_k)] \quad (2-)$$

(۴)

در آن چه در پی می‌آید برای نمایش سیستم‌های موجود در (۲-۱) و (۲-۳) از نماد گذاری زیر استفاده می‌نماییم:

$$y(n) = A(q)x(n) + N[x(n)] \quad (2-5)$$

در این جا q معرف عملگر تاخیر است و داریم

$$A(q) = \sum_{i=0}^N a_i q^{-i}$$

(۲-۶)

$$N[x(n)] = \sum_{k=2}^L \sum_{i_1=r}^N \dots \sum_{i_k=i_{k-1}}^N h_{i_1 i_2 \dots i_k} x(n-i_1) x(n-i_2) \dots x(n-i_k) \quad (2-7)$$

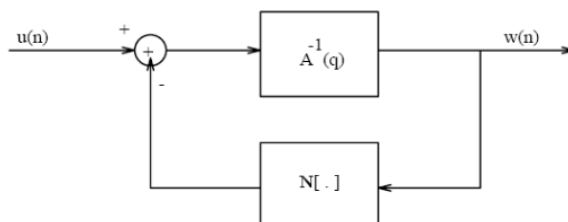
اگر $r=1$ ، (۲-۵) به (۲-۱) تبدیل می‌شود و اگر $r=0$ ، (۲-۵) به (۲-۴) تبدیل می‌شود. با این نماد گذاری می‌توانیم مشخصه ورودی - خروجی معکوس (۲-۵) را به صورت

$$A(q)w(n) = u(n) - N[w(n)] \quad (2-8)$$

نشان دهیم، که در آن $u(n)$ و $w(n)$ به ترتیب ورودی و خروجی سیستم معکوس هستند. می‌توانیم سیگنال خروجی را به این صورت نیز نمایش دهیم:

$$w(n) = A^{-1}(q)u(n) - A^{-1}(q)N[w(n)] \quad (2-9)$$

شکل (۳-۲) چند جمله ای موجود در (۹-۲) را به صورت گرافیکی نمایش می‌دهد. توجه کنید که می‌توانیم سیستم کلی را به صورت سیستمی فیدبک دار نشان دهیم که در فیدبک خود از بلوکی غیر خطی استفاده می‌کند و در بخش feed forward، معکوس بخش خطی سیستم اصلی را دارد.



شکل ۳-۲- نمایش گرافیکی سیستم معکوس سیستم توصیف شونده با سری Volterra

برای این که از پایداری سیستم معکوس (۵-۲) مطمئن باشیم باید بخش خطی سیستم غیر خطی مینیمم فاز باشد. در این صورت سیستم ارائه شده در (۹-۲) می‌تواند در ناحیه پایدار خود کار کند به شرطی که دامنه سیگنال ورودی به آن به اندازه کافی کوچک نگاه داشته شود. این که اندازه کافی برای کوچک بودن دامنه سیگنال ورودی چقدر است به صفرهای $A(q)$ و ضرایب موجود در $N[u(n)]$ بستگی پیدا می‌کند. بنابراین سیستم معکوس ارائه شده تنها در محدوده دامنه‌های سیگنال ورودی که آن را پایدار نگه می‌دارد می‌تواند عملکرد قابل قبولی داشته باشد.

در بسیاری از کاربردها علاقمندیم که سیستم مورد بحث را تنها در محدوده خاصی از فرکانس‌ها معکوس کنیم. مثلاً فرض کنید می‌دانیم که سیگنال ورودی به سیستم یک سیگنال تصویر با پهنای باند محدود است. چنین معکوس کننده ای می‌تواند از جابه جا کردن A^{-1} در شکل (۳-۲) با فیلتری که تنها در بازه فرکانسی دلخواه برابر A^{-1} و در بقیه فرکانس‌ها صفر است به دست آید. دو قضیه ای که در پی می‌آیند ساختار pre equalizer و post equalizer را برای کارکرد در ناحیه خاصی از طیف فرکانسی به دست می‌دهند:

قضیه (۱-۲): اگر سیگنال ورودی به سیستم غیر خطی تنها در باند خاصی مثل B غیر صفر باشد یک post equalizer برای سیستم (۵-۲) در بازه B به صورت زیر داده می‌شود:

$$w(n) = \tilde{A}^{-1}(q)u(n) - \tilde{A}^{-1}(q)N[w(n)]$$

(۱۰-۲)

که در اینجا $\tilde{A}^{-1}(q)$ سیستمی است که طیف آن در B برابر با اسپکتروم $A(q)$ ، و در بقیه فرکانس‌ها برابر صفر است.

اثبات: اثبات می‌کنیم سیستم معکوس ارائه شده در (۱۰-۲) در باند B خروجی را تولید می‌کند که برابر با ورودی سیستم غیر خطی است و در خارج از این باند خروجی برابر صفر تولید می‌کند. برای این منظور اتصال سیستم ارائه شده در (۵-۲) را با $\tilde{A}^{-1}(q)$ در نظر می‌گیریم و سپس سیستم حاصله را جبران سازی می‌کنیم. سری کردن سیستم ارائه شده در (۵-۲) با $\tilde{A}^{-1}(q)$ به وضوح منجر به صفر شدن خروجی در فرکانس‌های خارج از B می‌شود. سیستم حاصله دارای مشخصه ورودی-خروجی به شکل زیر خواهد بود:

$$z(n) = \tilde{A}^{-1}(q)A(q)x(n) + \tilde{A}^{-1}(q)N[x(n)] \quad (۱۱-۲)$$

از آنجا که $x(n)$ تنها در باند B دارای مولفه فرکانسی هست، معادله توصیف کننده سیستم مذکور برابر است با:

$$z(n) = x(n) + \tilde{A}^{-1}(q)N[x(n)] \quad (۱۲-۲)$$

که دارای سیستم post inverse به صورت زیر است:

$$w(n) = z(n) - \tilde{A}^{-1}(q)N[w(n)] \quad (۱۳-۲)$$

از آنجا که $z(n)$ نیز محدود به باند B است خروجی سیستم (۱۳-۲) نیز محدود به همین باند است. بنابراین cascade کردن $z(n) = \tilde{A}^{-1}(q)u(n)$ (که فرض می‌کنیم $u(n) = y(n)$) و سیستم (۲-۲) (۱۳) به وضوح معکوس (۵-۲) را در باند B می‌سازد. به سادگی می‌توان نشان داد که سیستم اخیر همان سیستم ارائه شده در (۱۰-۲) است.

قضیه (۲-۲): یک pre equalizer برای (۵-۲) در باند B به صورت زیر داده می‌شود:

$$w(n) = \tilde{A}^{-1}(q)u(n) - \tilde{A}^{-1}(q)N[w(n)]$$

که در اینجا $\tilde{A}^{-1}(q)$ همان تعریف قبلی را دارد.

اثبات این قضیه شبیه اثبات قضیه (۱-۲) است. توجه می‌کنیم که جبران سازهای ارائه شده در (۲-۱۰) و (۲-۱۳) برای سیستم توصیف شده در (۲-۱) و نه برای سیستم توصیف شده در (۲-۳) قابل کاربرد هستند.

۲-۶-۲- خطی سازهای ایده آل برای سیستمهای توصیف شونده با سری‌های ولترا

طریقه خطی سازی که در این جا ارائه می‌گردد همواره خروجی برابر قسمت خطی از سیستم غیر خطی (که قصد خطی سازی آن را داریم) را به دست می‌دهد. یک خطی ساز ایده آل سیستم غیر خطی داده شده را در تمام باند فرکانسی خطی می‌سازد.

قضیه (۲-۳): یک pre linearizer ایده آل برای سیستم (۲-۵) به صورت

$$w(n) = v(n) - A^{-1}(q)N[w(n)] \quad (2-14)$$

قابل نمایش است که در آن $v(n)$ و $w(n)$ به ترتیب بیان گر ورودی و خروجی آن هستند.

اثبات: با قرار دادن مشخصه ورودی - خروجی $u(n) = A(q)v(n)$ در (۲-۸) به دست می‌آوریم:

$$A(q)w(n) = A(q)v(n) - N[w(n)] \quad (2-15)$$

که دقیقاً همان سیستم (۲-۱۴) است. به طریقی مشابه می‌توانیم ثابت کنیم که یک post linearizer ایده آل برای (۲-۵) به صورت

$$w(n) = v(n) - N[A^{-1}(q)w(n)] \quad (2-16)$$

قابل تعریف است. این کار با قرار دادن $w(n) = A(q)x(n)$ و $v(n) = A(q)x(n) + N[x(n)]$ در رابطه فوق و بررسی صحت تساوی آن انجام می‌گیرد.

شکل (۲-۴) بلوک دیاگرام pre linearizer را نمایش می‌دهد. نمودار مشابهی می‌تواند برای post equalizer ترسیم شود. نتایج زیر می‌توانند به روشی شبیه روشی که برای اثبات قضیه (۲-۱) به کار رفت اثبات شوند.



شکل ۲-۴- بلوک دیاگرام pre linearizer سیستم توصیف شونده با سری Volterra

قضیه (۲-۴): اگر ورودی سیستم (۵-۲) محدود به پهنای باند B باشد و اگر $\tilde{A}^{-1}(q)$ تعریفی مشابه تعریف ارائه شده در قبل داشته باشد pre linearizer برای (۵-۲) در محدوده فرکانسی B دارای مشخصه ورودی - خروجی

$$w(n) = v(n) - \tilde{A}^{-1}(q)N[w(n)] \quad (۲-۸)$$

(۱۷)

خواهد بود.

قضیه (۲-۵): با تعریف قبلی برای $\tilde{A}^{-1}(q)$ ، post linearizer برای (۵-۲) در محدوده فرکانسی B دارای مشخصه ورودی - خروجی

$$w(n) = v(n) - N[\tilde{A}^{-1}(q)w(n)]$$

(۲-۱۸)

خواهد بود.

۲-۶-۳- جبران سازی مرتبه p برای سیستمهای توصیف شونده با سری‌های ولترا

جبران سازها و خطی سازهای ارائه شده در قسمت‌های قبلی شاید همیشه قابل به کار گیری نباشند. مثلاً هرگاه در (۵-۲) برابر با صفر باشد، جبران سازها و خطی سازهای ارائه شده دیگر فرم ساده ای ندارند. مشکلات مربوط به پایداری این سیستم‌ها نیز همواره وجود دارد. در آن چه که در پی می‌آید چند راهکار برای جبران سازی / خطی سازی مرتبه p سیستم‌های با مشخصه ورودی خروجی توصیف شده با سری ولترا ارائه می‌شود.

تعریف: یک جبران ساز / خطی ساز مرتبه p سیستمی است که قبل یا بعد از یک سیستم غیر خطی قرار می‌گیرد و باعث می‌شود سیستم کلی به دست آمده در بازه فرکانسی مورد نظر و در محدوده دامنه‌های مورد نظر از سیگنال‌های ورودی، دارای مشخصه ورودی - خروجی باشد که برابر است با مشخصه ورودی - خروجی یک سیستم متشکل از سیستمی همانی / خطی و یک سیستم غیر خطی

که مرتبه غیر خطی بودن آن برابر با p است. به این معنی که کوچکترین مرتبه جملات غیر خطی موجود در این مشخصه برابر p است.

قضیه (۶-۲): اگر سیستم (۸-۲) پایدار باشد دنباله سیستم‌های توصیف شده توسط

$$w_1(n) = A^{-1}(q)u(n) \quad (۱۹-۲)$$

$$w_p(n) = A^{-1}(q)u(n) - A^{-1}(q)N[w_{p-1}(n)] \quad (۲۰-۲)$$

به شرط میل دادن p به بی نهایت به سیستم (۸-۲) میل می‌کند. در ضمن (۲۰-۲) سیستم معکوس مرتبه p برای (۵-۲) است.

اثبات: کار را با استقرا انجام می‌دهیم. خروجی (۲۰-۲) به شرط داشتن ورودی برابر با (۵-۲) را به دست می‌آوریم. اگر $u(n) = y(n)$ ، برای $p=1$ رابطه ورودی - خروجی

$$\begin{aligned} w_1(n) &= x(n) + A^{-1}(q)N[x(n)] \\ &= x(n) + T_1(n) \end{aligned} \quad (۲۱-۲)$$

را داریم که در آن $T_1(n)$ یک سیستم با مرتبه بزرگتر از ۱ است. این موضوع قضیه را برای $p=1$ اثبات می‌کند. حال فرض کنید $w_p(n)$ خروجی یک معکوس مرتبه p است. یعنی

$$w_p(n) = x(n) + T_p(n) \quad (۲۲-۲)$$

که در آن $T_p(n)$ یک سیستم با مرتبه بزرگتر از p است. می‌خواهیم ثابت کنیم که سیستم توصیف شده توسط

$$w_{p+1}(n) = A^{-1}(q)u(n) - A^{-1}(q)N[w_p(n)] \quad (۲۳-۲)$$

یک معکوس با مرتبه $p+1$ برای (۵-۲) است. اگر (۵-۲) را در (۲۲-۲) و (۲۳-۲) قرار دهیم داریم

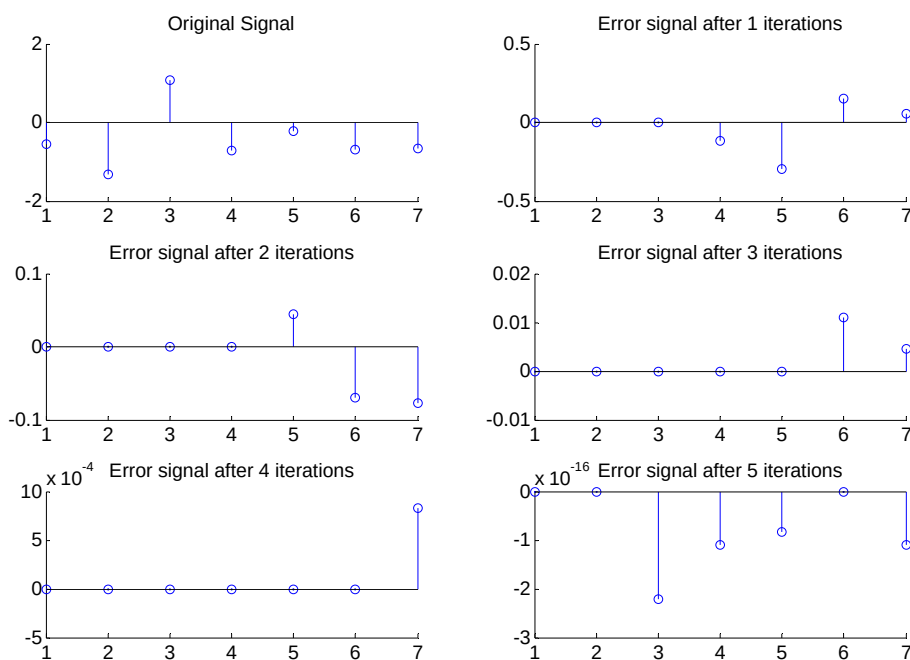
$$w_{p+1}(n) = x(n) + A^{-1}(q)N[x(n)] - A^{-1}(q)N[x(n) + T_p(n)]$$

$$= x(n) + A^{-1}(q)N[x(n)] - A^{-1}(q)N[x(n)] + T_{p+1}(n) \quad (24-2)$$

که در آن T_{p+1} یک سیستم با درجه بزرگتر از $p+1$ است. توجه کنید که این مطلب که $N[.]$ سیستمی با مرتبه بزرگتر از ۱ است را هم در نظر گرفته ایم. چون سیستم (۸-۲) معکوس سیستم (۲-۲) است و ما ثابت کردیم که دنباله تعریف شده با (۱۹-۲) و (۲۰-۲) معکوس‌های مرتبه p را برای (۵-۲) ایجاد می‌کنند، به شرط پایداری (۸-۲)، (۱۹-۲) و (۲۰-۲) به (۸-۲) همگرا می‌گردند.

نکته قابل توجه این روش معکوس سازی این است که همواره قابل پیاده سازی است در صورتی که این مطلب در مورد سیستم ارائه شده در (۳-۲) لزوماً درست نیست زیرا سیستم (۳-۲) ممکن است دچار مشکلات مربوط به پایداری باشد. تعمیم قضیه مربوط به جبران ساز / خطی سازهای مرتبه p به موارد ذکر شده در قضیه‌های (۱-۲) تا (۴-۲) به سادگی قابل انجام است.

به عنوان یک مثال از این روش معکوس سازی تقریبی سیستم‌های توصیف شونده با سری‌های ولترا، خطای تقریب‌های متوالی از ورودی سیستمی با معادله توصیف کننده به صورت $y[n] = x[n] + 0.7x[n-1]x[n-2]$ را از روی خروجی آن و با استفاده از (۱۹-۲) و (۲۰-۲) به دست می‌آوریم.



شکل ۲-۵- سیگنالهای خطای متوالی در بازسازی یک سیگنال از روی نسخه عبوری آن از سیستم ولترا

ملاحظه می‌کنیم که تقریب‌های متوالی به سرعت به سیگنال اصلی همگرا می‌شوند.

۲-۷- رگولاتورهای دیگر

روش‌های دیگری نیز برای رگولاسیون موجودند که از جمله آنها استفاده از فیلتر کالمن است. از روش‌های دیگری نیز که برای حذف نویز در سیستم‌های تفکیک مودی استفاده می‌شود می‌توان به عنوان رگولاتور استفاده نمود. یک نمونه از این روش‌ها تبدیل موجک است. هر گونه روش تخمین را نیز می‌توان به عنوان یک رگولاتور قلمداد کرد.

۲-۸- خلاصه

در این فصل مروری بر سیستم‌های معکوس و روش‌های رگولاسیون انجام شد، و نحوه خطی سازی سیستم‌های توصیف شونده با سری‌های ولترا بررسی گردید. در صورتی که هدف معکوس سازی و بعد تفکیک مودی تعمیم یافته یک سیستم توصیف شده با سری ولترا با استفاده از روشهای مورد بحث در این گزارش باشد، می‌توان ابتدا آن را به وسیله متدهای ارائه شده در این فصل تا مرتبه دلخواه خطی کرد و سپس وارد پروسه تفکیک مودی تعمیم یافته نمود. لذا این روش خطی سازی می‌تواند به عنوان یک متد رگولاسیون برای فرآیند معکوس سازی مورد مطالعه در این طرح پژوهشی تعبیر شود.

نظر به آن که برخی روشهای معکوس سازی برای تفکیک مودی تعمیم یافته مورد مطالعه در این نوشتار روشهایی لایه ای و مبتنی بر محاسبات عددی هستند، فصل بعد به مطالعه در مورد تکنیک‌های عددی در تفکیک مودی و معکوس سازی که با روشهای مورد نظر ما ارتباط نزدیکی دارند اختصاص یافته است.

فصل سوم

روش‌های عددی در معکوس‌سازی سیستم‌ها

۳-۱- مقدمه ای بر روش‌های مبتنی بر محاسبات عددی مورد استفاده در معکوس سازی سیستم

روش‌های معکوس سازی در مباحث مربوط به پردازش سیگنال از جایگاه ویژه ای برخوردارند. بسیاری از مسائلی که حل تحلیلی و مستقیم شناخته شده ای ندارند را می‌توان با استفاده از روش‌های معکوس سازی به صورت دقیق و یا با تقریب‌های بسیار خوبی حل کرد. دسته دیگری از مسائل نیز وجود دارند که علی‌رغم وجود حل مستقیم برای آنها به دلیل این که سیستم‌های مورد نیاز برای پیاده سازی این روش‌ها بسیار حساس و در مواردی ناپایدار هستند، استفاده از روش‌های معکوس سازی را می‌طلبند. در این جا مروری اجمالی بر برخی از مهم ترین روش‌های معکوس سازی مورد استفاده در تفکیک مودی تعمیم یافته که مرتبط با بحث اصلی ما هستند خواهیم داشت تا زمینه برای بحث اصلی مطرح شده در این گزارش هموار، و نیز امکان مقایسه بین روش‌های متفاوت با روشهای پیشنهادی ما از نظر سرعت همگرایی، دقت و میزان حافظه مورد نیاز برای خواننده مهیا گردد.

دو نمونه از روش‌هایی که در بازسازی سیگنال‌هایی که از سیستم‌های خطی عبور کرده اند مورد استفاده قرار می‌گیرند روش‌های Jacobi و Gauss-Seidel هستند. مهمترین کاربرد این دو روش در معکوس سازی ماتریس‌ها است. از روش‌های BCG، steepest descent، و بسیاری دیگر از روش‌های دیگر نظر می‌توان نام برد.

۳-۲- روش قالب‌ها مناسب برای سیگنالهای تنک

در بسیاری از مسائل مربوط به معکوس سازی، با معکوس سازی سیستم‌هایی از روی دنباله ای از سیگنال مانند $\langle f, e_n \rangle$ بر روی اتمهای خاصی رو به رو هستیم. در نظریه ویولت می‌توانیم به نمونه‌های برداشته شده از یک سیگنال با پهنای باند محدود و یا مقادیر ضرایب ویولت یا تبدیل فوریه زمان کوتاه و معکوس سازی از روی این گونه اطلاعات به عنوان نمونه‌هایی از چنین مسائلی اشاره کنیم. در این گونه مسائل سوالی که مطرح است بازسازی و یا حداقل تقریب زدن f از روی این داده‌ها است. به طور اصولی می‌توان این گونه مسائل را در نظریه قالب‌ها که توسط Duffin و Schaffer در ارتباط با سری‌های فوریه غیرهارمونیک مطرح شد دنبال کرد. این نظریه با مطرح شدن نظریه ویولت استفاده از نتایج آن رونق گرفت.

با استفاده از روش‌های تسریع که در جبر خطی استفاده می‌گردد می‌توان تکنیک‌های مورد استفاده در نظریه قالب‌ها را سرعت بخشید. در بخش بعد به معرفی دقیق تر نظریه قالب و الگوریتم‌های

استاندارد آن می‌پردازیم. پس از آن در باره روش‌های تسریع متد قالب^{۱۹} با استفاده از چند جمله‌ای‌ها توضیح می‌دهیم. استفاده از این متدهای تسریع باعث کارایی بهتر روش‌های موجود در نظریه اصلی قالب می‌گردد.

۳-۲-۱- معرفی قالب‌ها و اثبات بازسازی سیگنال از طریق روش قالب‌ها

فرض کنید f سیگنالی در فضای هیلبرت H با ضرب داخلی تعریف شده ای مانند $\langle \cdot, \cdot \rangle$ است و $\{e_n \mid n \in \mathbb{Z}\}$ هم دنباله ای از پایه‌ها در این فضا هستند. نرم $\|f\|^2 = \langle f, f \rangle$ معمولاً به انرژی سیگنال ارتباط دارد و ضرب داخلی $\langle f, e_n \rangle$ هم با توجه به تعریف به دست می‌آید. برای انجام یک مسئله بازسازی فرض می‌شود که f به طور یکتا از روی مجموعه $\langle f, e_n \rangle$ ها به دست می‌آید. این مطلب باعث می‌گردد که تعریف زیر را مطرح نماییم:

تعریف: دنباله ای مثل $\{e_n \mid n \in \mathbb{Z}\}$ در فضای هیلبرت یک قالب نامیده می‌شود اگر دو عدد مثبت مثل A و B وجود داشته باشند به طوری که برای هر تابع در این فضا $f \in H$ داشته باشیم:

$$A\|f\|^2 \leq \sum_{n \in \mathbb{Z}} |\langle f, e_n \rangle|^2 \leq B\|f\|^2 \quad (۱-۳)$$

هر دو عدد ثابتی که به جای A و B قرار گیرند حدود قالب نامیده می‌شوند.

با روی کار آمدن تئوری ویولت نظریه قالب پر اهمیت تر شده است. در سال‌های اخیر محققان زیادی در زمینه ریاضیات و نیز مهندسی از این مبحث در کارهای خود بهره گیری کرده اند. هر سیگنالی مانند f در فضای توصیف شده H در بالا می‌تواند به طور یکتا توسط دنباله $\{e_n \mid n \in \mathbb{Z}\}$ توصیف گردد. در واقع برای این کار به یک فرآیند تقریب های متوالی نیاز است. برای این منظور سیستم قالب را به صورت زیر معرفی می‌نماییم:

$$Sf = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n \quad (۲-۳)$$

لم (۱-۳): (الگوریتم قالب): فرض کنید $\{e_n \mid n \in \mathbb{Z}\}$ یک قالب برای H با باندهای A و B باشد. در این صورت هر تابع در H می‌تواند از روی ضرایب $\langle f, e_n \rangle$ به صورت زیر باز سازی گردد:

¹⁹ Method of Frames

$$f_n = f_{n-1} + \frac{2}{A+B} S(f - f_{n-1})$$

(۳-۳)

و $f = \lim_{n \rightarrow \infty} f_n$ ، که در باره خطای آن می‌توان گفت:

$$\|f - f_n\| \leq \left(\frac{B-A}{B+A} \right)^n \|f\|$$

(۴-۳)

چون سیستم قالب تنها به اطلاعات مربوط به $\langle f, e_n \rangle$ نیاز دارد، لذا می‌توان از روی این ضرایب f را بازسازی نمود.

اثباتی/بتکاری: چون $\langle Sf, f \rangle = \sum_n |\langle f, e_n \rangle|^2$ با استفاده از (۱-۳) به دست می‌آوریم

$$-\frac{B-A}{B+A} \|f\|^2 \leq \left\langle \left(I - \frac{2}{A+B} S \right) f, f \right\rangle \leq \frac{B-A}{B+A} \|f\|^2 \quad (۵-۳)$$

یعنی برای سیستم نرم داریم

$$\left\| I - \frac{2}{A+B} S \right\| \leq \frac{B-A}{B+A} < 1$$

(۶-۳)

از (۳-۳) به دست می‌آید

$$\begin{aligned} f - f_n &= f - f_{n-1} - \frac{2}{A+B} S(f - f_{n-1}) = \left(I - \frac{2}{A+B} S \right) (f - f_{n-1}) = \\ &\dots = \left(I - \frac{2}{A+B} S \right)^n (f - f_0) \end{aligned}$$

(۷-۳)

بنابر این نامساوی مطرح شده در (۶-۳) به دست می‌دهد:

$$\|f - f_n\| \leq \left\| I - \frac{2}{A+B} S \right\|_{op}^n \|f - f_0\| \leq \left(\frac{B-A}{B+A} \right)^n \|f\|$$

(۸-۳)

که همان چیزی است که به دنبال آن بودیم.

با توجه به (۳-۳) با هر تخمین اولیه ای از f می‌توان روش تقریب‌های متوالی را انجام داد و به نتیجه رسید اما به طور طبیعی $f_0 = 0$ و $f_1 = 2(A+B)^{-1}Sf$ انتخاب می‌گردند.

نرخ همگرایی این روش تقریب‌های متوالی به طور خاص به تقریب‌هایی که برای A و B به کار می‌رود بستگی دارد. به دست آوردن مقدار دقیق این باندها معمولاً یک کار پیچیده ریاضی است. در نظریه ویولت برای برخی قالب‌ها تحلیل‌های دقیقی وجود دارد که از روی آن حدود قالب به طور دقیق به دست می‌آید ولی در زمینه‌های دیگر عموماً وضع این گونه نیست. به این دلیل معمولاً از نظریه قالب از یک پارامتر آرامش^{۲۰} مثل λ به صورت زیر استفاده می‌شود:

$$f_n = f_{n-1} + \lambda S(f - f_{n-1})$$

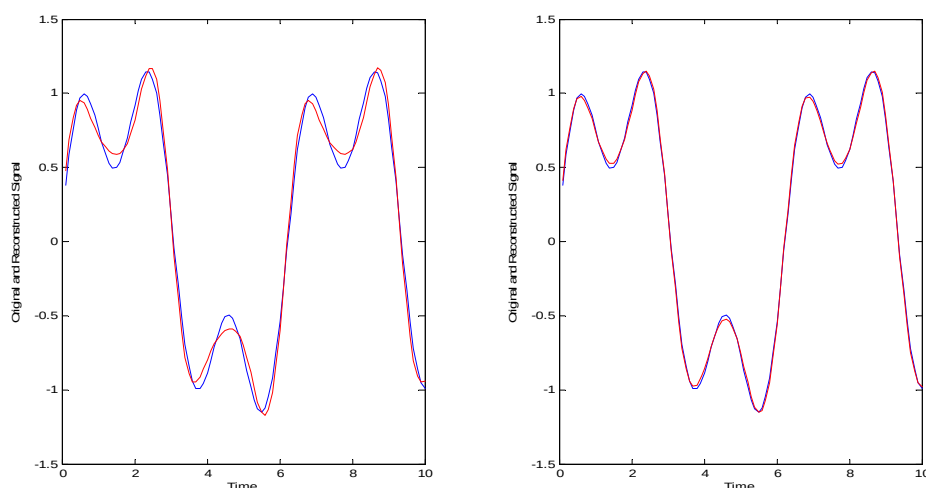
(۳-۹)

مشابه اثبات فوق به سادگی می‌توان ثابت کرد که اگر A و B معین باشند، انتخاب $\lambda = 2/(A+B)$ به صورت بهینه است و سریع‌ترین نرخ همگرایی را نتیجه می‌دهد. در عمل پارامتر λ به طور تجربی انتخاب می‌گردد و همین امر باعث می‌شود که الگوریتم قالب از سرعت کافی بر خوردار نباشد. اگر B معین باشد پارامتر $\lambda = 1/B$ می‌تواند یک پارامتر آرامش مورد قبول باشد، هر چند اگر A در دسترس نباشد نمی‌توان در مورد نرخ همگرایی به طور دقیق اظهار نظر کرد.

نظریه قالب و الگوریتم آن به طور طبیعی می‌تواند به فضاهاى دیگر (غیر از فضای هیلبرت) تعمیم داده شود. اما در فضاهاى دیگر روابط دیگر به صورت ارائه شده در بالا مطرح نیستند و تحلیل همگرایی الگوریتم پیچیده تر خواهد بود.

یک نمونه از کاربردهای نظریه قالب در بازسازی سیگنال‌های باند محدود از روی نمونه‌های آنها است. می‌توان قالبی ساخت که از سیگنال‌های باند محدود به همراه سیستم فیلترینگ ایده آل در لحظات نمونه برداری تشکیل شده است. با استفاده از نتایج ارائه شده در بالا توان گفت که سیگنال اولیه می‌تواند از روی نمونه‌های آن و با استفاده از رابطه تقریب‌های متوالی نظریه قالب به دست آید. شبیه‌سازی‌های زیر برای نمایش گرافیکی این موضوع انجام شده اند که سیگنال‌های اصلی و باز سازی شده را پس از دفعات مختلف استفاده از روش قالب نمایش می‌دهند.

²⁰ Relaxation Parameter



شکل ۳-۱- سیگنال اولیه به همراه بازسازی شده آن بعد از ۳ تکرار (چپ)- بعد از ۵ تکرار (راست)

۳-۲-۲- استفاده از رابطه تقریب‌های متوالی موجود در تئوری قالب با دید Successive Approximations

فرض کنید G یک سیستم خطی باشد که بر روی سیگنال x اثر کرده است. هدف آن است که نشان دهیم تحت شرایطی می‌توان x را از طریق فرمول تقریب‌های متوالی تئوری قالب بازسازی کرد بدون این که وارد مسائل مربوط به قالب بودن شویم. تقریب‌های متوالی y_n از x را چنین می‌سازیم ($y_0 = G(x)$):

$$y_n = y_{n-1} + \lambda(G(x) - G(y_{n-1}))$$

چون G خطی است می‌توان رابطه فوق را ساده کرد:

$$y_n = y_{n-1} + \lambda G(x - y_{n-1})$$

$$\Rightarrow e_n = y_n - x = (I - \lambda G)(e_{n-1})$$

اگر $I - \lambda G$ دارای این خاصیت باشد که بتوان نوشت $\|(I - \lambda G)x\| < \|I - \lambda G\| \|x\|$ ، می‌گوییم سیستم $E = I - \lambda G$ شرایط تقریب‌های متوالی^{۲۱} را ارضا می‌کند. در این صورت

$$\|e_n\| < \|E\| \|e_{n-1}\| < \dots < \|E^n\| \|e_0\|$$

²¹ Successive Approximations

اگر $\lim_{n \rightarrow \infty} E^n = \text{Zero Operator}$ آنگاه از رابطه بالا واضح است که $\lim_{n \rightarrow \infty} \|e_n\| = 0$. بنابر این یک شرط برای بازسازی سیگنال x با استفاده از فرمول بندی تئوری قالب آن است که λ بتواند چنان انتخاب شود که $E = I - \lambda G$ شرایط تقریب‌های متوالی را ارضا کند.

۳-۳- تسریع کننده‌های چند جمله‌ای مناسب برای سیگنالهای تنک

تا هنگامی که سرعت الگوریتم و میزان محاسبات اهمیتی نداشته باشند می‌توان از الگوریتم استاندارد قالب جهت بازسازی سیگنال‌ها استفاده کرد. اما برای قالب‌هایی با نسبت B/A بزرگ، الگوریتم نیاز به تعداد بسیار زیادی تکرار دارد تا بتواند بازسازی را به طریقی قابل قبول انجام دهد. برای مثال برای $B/A = 2000$ نرخ همگرایی برابر خواهد بود با 0.999 . در این حالت برای تضمین کردن $\|f - f_n\|/\|f\| \leq 0.01$ نیاز به انجام ۴۶۰ تکرار داریم و این از لحاظ عملی قابل انجام نیست. برای این که این مشکل بر طرف شود می‌توانیم از روش‌هایی که در جبر خطی شناخته شده هستند با انجام تغییراتی استفاده کنیم. ایده اصلی در این روش‌ها این است که در هر مرحله بهترین تقریب ممکن از f را بر حسب همه تقریب‌های قبلی آن، و نه تنها بر اساس آخرین تقریب به دست آمده به دست دهیم. یعنی:

$$h_n = \sum_{k=1}^n a_{nk} f_k \quad (10-3)$$

شرطی که باید لحاظ شود این است که

$$\sum_{k=1}^n a_{nk} = 1 \quad (11-3)$$

تا در هنگامی که $f_1 = \dots = f_n = f$ ، داشته باشیم $h_n = f$.

با استفاده از (۷-۳) و (۱۱-۳) خطا برابر خواهد بود با

$$f - h_n = \sum_{k=1}^n a_{nk} (f - f_k) = \sum_{k=1}^n a_{nk} \left(I - \frac{2}{A+B} S\right)^k (f - f_0) \quad (12-3)$$

و با تعریف $Q_n(x) = \sum_{k=1}^n x^k$ و $R = I - 2/(A+B)S$ داریم

$$f - h_n = Q_n(R)(f - f_0) \quad (۱۳-۳)$$

به این ترتیب خطا در هر مرحله تابعی چند جمله‌ای از خطای اولیه است. فرض کنید یک چند جمله‌ای پیدا شده است که می‌تواند خطای مرحله n ام را به اندازه کافی کوچک کند. مشکلی که وجود دارد این است که باید تقریب‌های مراحل قبلی محاسبه و در جایی ذخیره شده باشند تا بتوان از این چند جمله‌ای استفاده کرد. خوشبختانه روش‌هایی وجود دارند که می‌توانند بدون محاسبه مستقیم این‌ها و تنها با استفاده از بازگشت‌های سه ترمی^{۲۲} (که چند جمله‌ای مربوط به هر تکرار را بر حسب سه تا از چند جمله‌ای‌های قبلی به دست می‌دهند) این الگوریتم را پیاده‌سازی کنند. در ظاهر فرمول‌های تقریب‌های متوالی که این الگوریتم‌ها استفاده می‌کنند تمام چند جمله‌ای‌های بهینه قبلی دیده نمی‌شوند اما در به دست آوردن آنها اثر همگی چند جمله‌ای‌های قبلی شرکت داده شده‌اند. در این جا دو نمونه از این روش‌های تقریب‌های متوالی سرعت یافته را بررسی می‌کنیم.

۳-۴- تسریع به روش Chebyshev

این روش از چند جمله‌های چبیشف برای تخمین استفاده می‌کند. این روش برای مدت‌ها در مباحث مربوط به حل دستگاه‌های بزرگ خطی استفاده می‌شده است.

قضیه (۳-۱): (روش تسریع چبیشف): قالبی با سیستم S را در نظر بگیرید و قرار دهید $\rho = (B - A)/(B + A)$. در این صورت f می‌تواند با استفاده از $h_0 = 0$ و $h_1 = 2(A + B)^{-1}Sf$ و $\lambda_1 = 2$ به صورت زیر بازسازی شود:

$$h_n = \lambda_n(h_{n-1} - h_{n-2} + \frac{2}{A+B}S(f - h_{n-1})) + h_{n-2} \quad (۱۴-۳)$$

که در آن

$$\lambda_n = (1 - \frac{\rho^2}{4}\lambda_{n-1})^{-1} \quad (۱۵-۳)$$

اگر قرار دهیم $\sigma = \frac{\sqrt{B} - \sqrt{A}}{\sqrt{B} + \sqrt{A}}$ ، تقریبی که برای خطا به دست می‌آید به صورت زیر است.

$$\|f - h_n\| \leq \frac{2\sigma^n}{1 + \sigma^{2n}} \|f\|$$

(۱۶-۳)

پارامتر آرامش این الگوریتم به صورت وفقی تنظیم نمی‌شود.

می‌توان برای افزایش سرعت همگرایی از الگوریتم وفقی Conjugate Gradient استفاده کرد. این روش در سال‌های اخیر مقبول‌ترین روش برای حل دستگاه‌های خطی بزرگ محسوب می‌گردد. در این جا خلاصه‌ای از مراحل این الگوریتم را که قابل پیاده سازی روی هر گونه سیستم خطی است بیان می‌داریم.

۳-۵- Conjugate Gradient روش

قضیه (۲-۳): (الگوریتم تسریع کننده CG برای قالب‌ها): فرض کنید $\{e_n \mid n \in Z\}$ قالبی روی H با سیستم S باشد. سیگنال‌های h و r و p را با شروع از

$$h_0 = 0, \quad r_0, p_0 = Sf, \quad p_{-1} = 0 \quad (۱۷-۳)$$

به ترتیب زیر می‌سازیم:

$$h_{n+1} = h_n + \lambda_n p_n \quad (۱۸-۳)$$

$$r_{n+1} = r_n - \lambda_n S p_n \quad (۱۹-۳)$$

$$\lambda_n = \frac{\langle r_n, p_n \rangle}{\langle p_n, S p_n \rangle} \quad (۲۰-۳)$$

$$p_{n+1} = S p_n - \frac{\langle S p_n, S p_n \rangle}{\langle p_n, S p_n \rangle} p_n - \frac{\langle S p_n, S p_{n-1} \rangle}{\langle p_{n-1}, S p_{n-1} \rangle} p_{n-1} \quad (۲۱-۳)$$

اگر a و b به ترتیب کوچکترین و بزرگترین مقدار اسپکتروم سیستمی S را نشان دهند یعنی

$$a = \|S^{-1}\|_{op}^{-1} \text{ و } b = \|S\|_{op} \text{، خطای این روش از روش زیر تخمین زده می‌شود: } \sigma = \frac{\sqrt{b} - \sqrt{a}}{\sqrt{b} + \sqrt{a}}$$

$$\|f - h_n\|_S \leq \frac{2\sigma^n}{1 + \sigma^{2n}} \|f\|_S \quad (۲۲-۳)$$

تعداد iterationهای لازم برای رسیدن به خطای نسبی از پیش تعیین شده ای مثل $\|f - f_n\|/\|f\| < \varepsilon$ به صورتی که در این جا شرح داده می‌شود محاسبه می‌گردد. فرض کنید $k = B/A$ عدد بزرگی است وضرب‌های همگرایی در الگوریتم استاندارد و تسریع شده قالب به ترتیب برابر $\rho = \frac{B-A}{B+A}$ و $\sigma = \frac{\sqrt{B}-\sqrt{A}}{\sqrt{B}+\sqrt{A}}$ باشند. در این صورت در الگوریتم استاندارد قالب داریم

$$\|f - f_n\|/\|f\| < \varepsilon \Rightarrow \rho^n < \varepsilon \Rightarrow n_{frame} > \frac{\log \frac{1}{\varepsilon}}{\log \frac{1}{\rho}} \approx \frac{k-1}{2} \log \frac{1}{\varepsilon} \quad (23-3)$$

این به این دلیل است که $\log(1/\rho) = \log \frac{k+1}{k-1} = \log(1 + \frac{2}{k-1}) \approx \frac{2}{k-1}$.

در الگوریتم تسریع شده، $\|f - f_n\|/\|f\| < \varepsilon$ لازم می‌دارد که $2\sigma^n(1 + \sigma^{2n})^{-1} < \varepsilon$ ، یا به طور غیر دقیق تر $\sigma^n < \varepsilon/2$ ، که نتیجه می‌دهد:

$$n_{acc} > \frac{\log \frac{2}{\varepsilon}}{\log \frac{1}{\sigma}} \approx \frac{\sqrt{k}-1}{2} (\log 2 + \log \frac{1}{\varepsilon}) \quad (24-3)$$

به این ترتیب ملاحظه می‌شود که در حالاتی که k بزرگ است استفاده از روش‌های تسریع کننده کاملاً مفید است.

روش چیشف بهترین تخمین ممکن به صورت چند جمله ای شامل تخمین‌های قبلی را به دست می‌دهد ولی نیاز به دانستن حدود قالب دارد. روش CG نیازی به دانستن این حدود ندارد.

یکی از مهمترین کاربردهای روش Conjugate Gradient در حل معادلات ماتریسی است که در این جا به بررسی آن می‌پردازیم. به عبارت دیگر در اینجا مسئله معکوس کردن ماتریس‌ها توسط این متد را مورد اثبات قرار می‌دهیم.

۳-۵-۱ – استفاده از روش Conjugate Gradient در حل معادلات ماتریسی

در زمینه‌های مرتبط با پردازش سیگنال در موارد متعددی لازم می‌شود که معادله ای ماتریسی

شبه به

$$Ax = b \quad (25-3)$$

حل شود که در آن A ماتریسی $n \times n$ و b برداری $n \times 1$ است. به دنبال یافتن روش سریعی برای حل آن هستیم به نحوی که الگوریتم حاصله را بتوان در زمان حقیقی یا $real\ time$ پیاده سازی نمود.

در روش CG که آن را روش حذف تصاویر خطا نیز می‌نامند، به الگوریتم مناسبی با ویژگی‌های اشاره شده در بالا دست می‌یابیم. فرمول بندی این روش در بخش قبلی بیان گردیده است. در این بخش سعی داریم ایده استفاده شده در این الگوریتم را در قالب مثال وارون سازی ماتریس‌ها شرح دهیم. ایده اصلی این الگوریتم به این صورت است:

K را زیر فضایی از R^n بگیرید. فرض کنید در مرحله m ام از iteration موجود در الگوریتم هستیم. تقریب m ام از x را با اعمال m شرط روی زیر فضای K (به طوری که در ادامه خواهیم دید) به دست می‌آوریم. یک روش برای اعمال این m شرط قرار دادن m شرط تعامد روی زیر فضای K است. به بیان دقیق تر این شرط را می‌گذاریم که $b - Ax_m$ به m عضو از اعضای K عمود باشد. لذا x_m بهترین تخمین خطی از x با استفاده از اعضای K خواهد بود. در بخش‌های بعدی این ایده قدم به قدم توضیح داده خواهد شد.

۳-۵-۱-۱- تعاریف

در این بخش چند تعریف برای مفاهیمی که در ادامه بحث به کار گرفته خواهند شد ارائه می‌شود.

- ماتریس نیمه مثلثی: ماتریسی را نیمه مثلثی می‌نامیم که برای هر i و j به نحوی که $i > j + 1$ ، درایه a_{ij} آن صفر باشد.
- ماتریس سه قطری: ماتریسی است که تنها عناصر غیر صفر آن a_{ii} ، $a_{i(i+1)}$ و $a_{(i+1)i}$ باشند.
- چند جمله ای مینیمال $P_A(V) : A$ را یک ماتریس $n \times n$ و V را یک بردار $n \times 1$ در نظر بگیرید. چند جمله ای $P(t)$ به صورت $P(t) = a_0 + a_1 t + \dots + a_m t^m$ را چند جمله ای مینیمال V نسبت به ماتریس A می‌نامیم و با $P_A(V)$ نشان می‌دهیم اگر $P(A).V = 0$ ، و هیچ چند جمله ای با درجه کمتر از m وجود نداشته باشد که این خاصیت برقرار باشد. در این صورت درجه V را نسبت به A برابر m می‌گیریم.

۳-۵-۱-۲- ایده اصلی

فرض کنید A یک ماتریس $n \times n$ ، و K یک زیر فضای m بعدی از R^n باشد. تقریب x تا لحظه $m-1$ را x_{m-1} می‌نامیم. روش حذف تصاویر خطا پروسه ای است که جواب تقریبی دیگری مثل x_m برای $(25-3)$ پیدا می‌کند با این شرط که $x_m \in K_m$ و همچنین بردار خطای باقی مانده بعد از تبدیل ماتریسی بر K_m عمود باشد. یعنی

$$\text{find } x_m \in K_m \text{ such that } (b - Ax_m) \perp K_{m-1} \quad (26-3)$$

اگر علاقمند باشیم که تخمین ما راجع به x در بهبود تقریب‌های بعدی تاثیر داده شوند باید تخمین خود در هر مرحله را به جای این که روی زیر فضای K_{m-1} انجام دهیم روی $x_{m-1} + K_{m-1}$ انجام دهیم. بنا بر این عبارات بالا اندکی تغییر می‌یابند.

$$\text{Find } x_m \in (x_{m-1} + K_{m-1}) \text{ such that } (b - Ax_m) \perp (x_{m-1} + K_{m-1}) \quad (27-3)$$

توجه کنید که اگر x_m به فرم $x_m = x_{m-1} + \delta$ نوشته شود، باقی مانده مرحله قبل را این گونه تعریف کنیم که $r_{m-1} = b - Ax_{m-1}$ و زیر فضای K_m جدید را برابر $x_{m-1} + K_{m-1}$ در نظر بگیریم شرط $(27-3)$ تبدیل می‌شود به

$$b - A(x_{m-1} + \delta) \perp K_{m-1} \quad (28-3)$$

$$(r_{m-1} - A\delta) \perp K_{m-1} \quad (29-3)$$

به بیان دیگر جواب تقریبی می‌تواند به صورت $x_m = x_{m-1} + \delta$ تعریف شود، که برای هر $w \in K_m$ داریم $(r_{m-1} - A\delta, w) = 0$. در این جا منظور از (x, y) ضرب داخلی بردار x و بردار y است.

مراحل ذکر شده در بالا یک قدم در روش حذف تصاویر خطا است. می‌توان در هر مرحله K را عوض کرد یا آن را گسترش داد.

نتیجه بلا فاصله ای که از $(29-3)$ به دست می‌آید این است که $\|r_m\| < \|r_{m-1}\|$ ، زیرا r_m از تفاضل r_{m-1} و تصویر آن روی یک زیر فضا به دست آمده است و بنابراین نرم کوچکتری نسبت به نرم r_{m-1} دارد.

۳-۱-۵-۳- معرفی یک زیر فضای خاص از R^n

برای این که زیر فضای K را در هر مرحله از الگوریتم به طور ضمنی بسازیم و لازم نباشد که برای ساختن آن محاسبات مجزایی انجام دهیم، K را به طور خاص به این روش می‌سازیم که در مرحله m ام:

$$K_m(A, r_0) = \text{span}\langle r_0, Ar_0, A^2r_0, \dots, A^{m-1}r_0 \rangle \quad (۳۰-۳)$$

که همان گونه که قبلاً تعریف شد. $r_0 = b - Ax_0$

اگر بخواهیم از منظر تئوری تقریب به قضیه نگاه کنیم، با K_m تعریف شده در بالا، تقریبی که در مرحله m ام برای x به دست می‌آید به فرم زیر است:

$$A^{-1}b \approx x_m = x_0 + q_{m-1}(A)r_0 \quad (۳)$$

(۳۱)

که q_{m-1} یک چند جمله ای با درجه $m-1$ است.

اولین ساده سازی که به ذهن می‌رسد این است که x_0 را برابر صفر قرار دهیم و در این صورت

$$A^{-1}b \approx x_m = q_{m-1}(A)r_0 \quad (۳)$$

(۳۲)

بدیهی است که روشی که برای ساختن $K_m(A, r_0)$ اتخاذ شده است باعث می‌شود که بعد $K_m(A, r_0)$ در هر مرحله یک واحد اضافه شود. و این دقیقاً همان ویژگی است که باعث می‌شود خطا در تقریب در هر مرحله کمتر از خطا در مرحله قبلی باشد. یکی از خواص $K_m(A, r_0)$ این است که فضایی است که می‌توان هر بردار به شکل $y = P(A)r_0$ با درجه کمتر یا مساوی m را در آن بسط داد.

۳-۵-۱-۴- ساختار اصلی الگوریتم حذف تصویرهای خطا

با توجه به مطالبی که عنوان شدند ساختار اصلی الگوریتم حذف تصویرهای خطا به صورت زیر معرفی می‌گردد. اگر در این جا r_0 را با v_1 ، r_1 را با v_2 ، و... نشان دهیم داریم:

Algorithm 1

1-Choose a vector v_1 such that $|v_1| = 1$.

2-For $j=1,2,\dots,m$ Do

3- Compute $h_{ij} = (Av_i, v_j)$ for $i=1,2,\dots,j$

4- Compute $w_j = Av_j - \sum_{i=1}^j h_{ij}v_i$

5- $h_{j+1,j} = \|w_j\|$

6-If $h_{j+1,j} = 0$, stop.

7- $v_{j+1} = w_j / h_{j+1,j}$

در مرحله j ام، الگوریتم v_j را در A ضرب می‌کند، بردار به دست آمده را روی همه v_i های قبلی ($i < j$) تصویر می‌کند، و سپس آن را نرمالیزه می‌نماید. بنابر این یک عملیات معمولی گرام اشمیت انجام می‌دهد. الگوریتم در صورتی خاتمه می‌یابد که w_j محاسبه شده در خط چهارم آن مساوی با صفر شود. در بخش‌های بعدی این حالت را بیشتر مورد مطالعه قرار می‌دهیم. در این جا چند خاصیت این الگوریتم بررسی می‌شود.

قضیه (۳-۳) : فرض کنید که الگوریتم قبل از مرحله m ام خاتمه نیافته باشد. در این صورت v_1 و v_2 و ... و v_m یک پایه متعامد یکه را در R^n تشکیل می‌دهند که فضای زیر را می‌سازند:

$$K_m = \text{span}\langle v_1, Av_1, A^2v_1, \dots, A^{m-1}v_1 \rangle$$

اثبات: بردارهای v_j $j=1,\dots,m$ طبق روش ساخت متعامد یکه هستند. این که آنها می‌توانند K_m را span کنند از این واقعیت نتیجه می‌شود که هر کدام از v_j ها به فرم $q_{j-1}(A)v_1$ هستند که q_{j-1} یک چند جمله ای با درجه $j-1$ است. این ادعا می‌تواند همانند زیر با استقرا نشان داده شود.

برای $j=1$ اثبات بدیهی است. زیرا برای $j=1$ ، $v_1 = q_0(A)v_1$ که $q_0(t)=1$. فرض کنید که نتیجه برای هر عدد کمتر یا مساوی j برقرار باشد و v_{j+1} را در نظر بگیرید. داریم:

$$h_{j+1,j}v_{j+1} = Av_j - \sum_{i=1}^j h_{ij}v_i = Aq_{j-1}(A)v_1 - \sum_{i=1}^j h_{ij}q_{j-1}(A)v_1$$

(۳۳-۳)

که نشان می‌دهد که v_{j+1} نیز می‌تواند به صورت $q_j(A)v_1$ نمایش داده شود که q_j یک چند جمله ای از درجه j است. به این ترتیب اثبات کامل می‌شود.

قضیه (۳-۴): فرض کنید V_m ماتریس $n \times m$ ای را نشان دهد که ستون‌های آن v_1 تا v_m باشند. $\bar{H}_m$ را ماتریس $(m+1) \times m$ ای در نظر بگیرید که نیمه مثلثی است و عناصر غیر صفر آن توسط الگوریتم ۱ به دست می‌آیند، و H_m را ماتریس حاصل از حذف سطر آخر $\bar{H}_m$ در نظر بگیرید. بنابر این روابط زیر را داریم:

$$AV_m = V_m H_m + w_m e_m^T \quad e_m = [000 \dots 1] \quad (\text{size of } e_m = 1 \times m) \quad (34-3)$$

$$= V_{m+1} \bar{H}_m \quad (3)$$

$$V_m^T AV_m = H_m \quad (35)$$

$$(36-3)$$

اثبات: (۳-۳) از تساوی زیر به دست می‌آید که آن نیز از خطوط ۴ و ۵ و الگوریتم ۱ به دست می‌آید.

$$Av_j = \sum_{i=1}^{j+1} h_{ij} v_i \quad j = 1, \dots, m \quad (37-3)$$

(۳-۴) بیان دیگر (۳-۳) به صورت ماتریسی است. (۳-۳۶) نیز از ضرب طرفین (۳-۳۴) در V_m^T و استفاده از متعامد یک‌به‌یک بودن $v_1, \dots, v_m$ به دست می‌آید ($V_m^T V_m = I_{m \times m}$).

همان گونه که قبلاً نیز بیان شد الگوریتم ۱ ممکن است در مرحله j ام چنانچه w_j برابر صفر شود متوقف شود. اکنون شرایطی را بررسی می‌کنیم که تحت آن این اتفاق می‌افتد.

قضیه (۳-۵): الگوریتم (۱) در مرحله j ام متوقف می‌شود، یعنی $h_{j+1,j} = 0$ ، اگر و تنها اگر چند جمله ای مینیمال v_1 از درجه j باشد.

اثبات: اگر درجه چند جمله ای v_1 برابر j باشد باید w_j برابر با صفر باشد در غیر این صورت v_{j+1} قابل تعریف شدن است و بنابر این $K_{j+1}(A, v_1)$ از بعد $j+1$ است. اما در این صورت μ ، درجه چند جمله ای مینیمال v_1 نسبت به A حداقل $j+1$ است و این تناقض است.

برای اثبات عکس قضیه فرض کنید $w_j = 0$. در آن صورت درجه چند جمله ای مینیمال v_1 نسبت به A چنان است که $j \leq \mu$. از طرفی غیر ممکن است که $\mu < j$ ، زیرا در این صورت طبق

قسمت اول اثبات باید w_μ صفر باشد و این بدین معنی است که الگوریتم باید در مرحله با شماره کوچکتر از j یعنی μ متوقف می‌شد. بنابر این قضیه (۳-۵) به طور کامل اثبات شد. اگر زمانی الگوریتم در مرحله j ام تمام شود به این معنی است که x به طور دقیق به دست آورده شده است.

۳-۵-۱-۵-۱ الگوریتم (۲) حذف تصاویر خطا (برای حالتی که A متقارن است).

اگر A متقارن باشد ماتریس H_m تبدیل به یک ماتریس سه قطری می‌شود، و این باعث می‌گردد که متعامد سازی به روش گرام اشمیت انجام شده در الگوریتم (۱) در هر مرحله تنها به دو بردار قبلی بستگی داشته باشد. یعنی برای ساختن v_{j+1} تنها به v_j و v_{j-1} نیاز داشته باشیم.

قضیه (۳-۶): فرض کنید الگوریتم (۱) روی یک ماتریس حقیقی A که متقارن است انجام شده باشد. در این صورت ضرایب h_{ij} که به دست می‌آیند چنان اند که:

$$h_{ij} = 0 \quad \text{for} \quad 1 \leq i \leq j-1 \quad (3-38)$$

$$h_{j,j+1} = h_{j+1,j} \quad \text{for} \quad j = 1, 2, \dots, m \quad (3-39)$$

به عبارت دیگر H_m یک ماتریس سه قطری و متقارن است.

اثبات: اثبات این قضیه نتیجه بلافاصله این حقیقت است که $H_m = V_m^T A V_m$ یک ماتریس متقارن است که در ضمن مثلثی هم هست.

اگر نماد گذاری‌هایمان را تغییر دهیم به نحوی که $\alpha_j = h_{jj}$ و $\beta_j = h_{j-1,j}$ آنگاه داریم:

$$H_m = \begin{bmatrix} \alpha_1 & \beta_1 & 0 & \dots \\ \beta_1 & \alpha_2 & \beta_2 & \dots \\ 0 & \beta_2 & \ddots & \\ 0 & 0 & 0 & \end{bmatrix} \quad (3-40)$$

لذا الگوریتم (۲) حذف تصاویر خطا با اندکی تغییر از روی الگوریتم (۱) ساخته می‌شود.

Algorithm 2

1-Choose an initial vector v_1 of norm 1, set $\beta_1 = 0$ and $v_0 = 0$.

2- For $j=1,2,\dots,m$ Do

$$3- w_j = Av_j - \beta_j v_{j-1}$$

$$4- \alpha_j = (w_j, v_j)$$

$$5- w_j = w_j - \alpha_j v_j$$

$$6- \beta_{j+1} = |w_j| \quad \text{If } \beta_{j+1} = 0, \text{ then stop.}$$

$$7- v_{j+1} = w_j / \beta_{j+1}$$

توجه: هر چند به نظر عجیب می‌آید که این الگوریتم تضمین کند که v_j ‌های به دست آمده بر هم عمود باشند اما چنان که اثبات‌های موجود در قسمت‌های قبلی نشان دادند این خاصیت وجود دارد.

۳-۵-۱-۶- الگوریتم (۳) الگوریتم نهایی و بهینه CG

اکنون به ابزارهای لازم مجهز هستیم تا بتوانیم مسئله اصلی را که همان به دست آوردن x از رابطه $Ax = b$ است حل کنیم. در این جا ابتدا با ضرب A در A^T به ماتریس جدیدی می‌رسیم که متقارن است و آن را A می‌نامیم. پس می‌توانیم از الگوریتم (۲) استفاده نماییم.

در بخش‌های قبلی ثابت کرده بودیم که r_m برابر است با $r(A) \times r_0$ که $r(t)$ یک چند جمله‌ای بر حسب t است. از ترکیب این خاصیت با این واقعیت که الگوریتم (۲) $r_{m+1}(t)$ را تنها از روی $r_m(t)$ و $r_{m-1}(t)$ می‌سازد به رابطه زیر می‌رسیم:

$$r_{m+1}(t) = \rho_m(r_m(t) - \gamma_m t r_m(t)) + \delta_m r_{m-1}(t)$$

(۴۱-۳)

باید داشته باشیم $r_m(0) = 1$. با اعمال این شرط به (۴۱-۳) به دست می‌آوریم که $\delta_m = 1 - \rho_m$. لذا

$$r_{m+1}(t) = \rho_m(r_m(t) - \gamma_m t r_m(t)) + (1 - \rho_m)r_{m-1}(t)$$

(۴۲-۳)

و یا در حوزه ماتریسی

$$r_{m+1} = \rho_m(r_m - \gamma_m A r_m) + (1 - \rho_m)r_{m-1}$$

(۴۳-۳)

باید $(r_{m+1}, r_m) = 0$. با اعمال این شرط به (۴۳-۳) و استفاده از تعامد r_m و r_{m-1} به دست می‌آوریم:

$$\gamma_m = \frac{(r_m, r_m)}{(Ar_m, r_m)}$$

(۴۴-۳)

همین طور باید $(r_{m+1}, r_{m-1}) = 0$. با اعمال این شرط به (۴۳-۳) و انجام برخی ساده سازی‌های جبری نتیجه می‌گیریم:

$$\rho_m = \left[1 - \frac{\gamma_m}{\gamma_{m-1}} \cdot \frac{(r_m, r_m)}{(r_{m-1}, r_{m-1})} \cdot \frac{1}{\rho_{m-1}} \right]^{-1}$$

(۴۵-۳)

تقریب‌های متوالی برای x می‌توانند از روی تقریب‌های متوالی برای r به دست آیند. از مطالب بالا واضح است که می‌توان چند جمله‌ای‌های $s_1, s_2, \dots$ از درجات برابر با اندیششان را چنان ساخت که:

$$r_m(t) = 1 - ts_{m-1}(t) \quad \Rightarrow \quad s_m(t) = \frac{1 - r_{m+1}(t)}{t} =$$

(۴۶-۳)

$$\begin{aligned} &= \rho_m \left(\frac{1 - r_m(t)}{t} + \gamma_m r_m(t) \right) + (1 - \rho_m) \frac{1 - r_{m-1}(t)}{t} \\ &= \rho_m (s_{m-1}(t) + \gamma_m r_m(t)) + (1 - \rho_m) s_{m-2}(t) \end{aligned}$$

نتیجه بلافاصله $r_m(t) = 1 - ts_{m-1}(t)$ این است که

(۴۷-۳)

$$r_m(t) = 1 - ts_{m-1}(t) \Rightarrow b - Ax_m = b - Ax_0 - As_{m-1}(A)r_0 \Rightarrow x_m = x_0 + s_{m-1}(A)r_0$$

ترکیب رابطه (۴۶-۳) و رابطه (۴۷-۳) رابطه تقریب‌های متوالی زیر را برای یافتن تقریب‌های متوالی x (و نه خطاهای متوالی) به دست می‌دهد:

$$x_{m+1} = \rho_m (x_m + \gamma_m r_m) + (1 - \rho_m) x_{m-1}$$

(۴۸-۳)

آن چه برای شروع عملیات تقریب‌های متوالی باقی مانده است تنها معین کردن x_0 و x_1 است. x_0 را می‌توان برداری دلخواه در نظر گرفت. بردار x_1 باید برداری به شکل زیر باشد:

$$x_1 = x_0 + \gamma_0 r_0$$

(۴۹-۳)

نتیجه (۴۸-۳) و (۴۹-۳) این است که باید قرار دهیم $\rho_0 = 1$ و $x_{-1} = 0$. خلاصه بحث انجام شده منجر به الگوریتم شماره (۳) CG به صورت زیر می‌شود:

Algorithm 3Set $x_0 = 0$ 1- Compute $r_0 = b - Ax_0$. Set $x_{-1} = 0$ and $\rho_0 = 1$.2- For $j=1,2,\dots$, until convergence Do3- Compute Ar_j and $\gamma_j = \frac{(r_j, r_j)}{(Ar_j, r_j)}$ 4- If $j > 0$ compute $\rho_j = \left[1 - \frac{\gamma_j}{\gamma_{j-1}} \cdot \frac{(r_j, r_j)}{(r_{j-1}, r_{j-1})} \cdot \frac{1}{\rho_{j-1}} \right]^{-1}$ 5- Compute $x_{j+1} = \rho_j(x_j + \gamma_j r_j) + (1 - \rho_j)x_{j-1}$ 6- Compute $r_{j+1} = \rho_j(r_j - \gamma_j Ar_j) + (1 - \rho_j)r_{j-1}$

شبیه‌سازی‌ها نشان می‌دهند که الگوریتم (۳) برای به دست آوردن جواب یک معادله ماتریسی حدود ربع زمانی را که MatLab برای محاسبه لازم دارد نیاز دارد. چنان که عنوان شد، تعمیم این روش به فضاهای غیرماتریسی نسبتاً به سادگی امکان پذیر است. کافی است ضرب داخلی استفاده شده در الگوریتم‌های ۱ و ۲ را ضرب داخلی تعریف شده روی فضای مورد نظر در نظر گرفته و از الگوریتم‌های فوق استفاده نماییم.

الگوریتم بعدی، روش تصویر کردن‌های متوالی بر روی مجموعه‌های محدب است. به منظور آشنایی خواننده با این مفهوم، در این جا خلاصه ای از این متد را بیان می‌نمایم.

۳-۶- تصویر کردن بر روی مجموعه‌های محدب^{۲۳} مناسب برای سیگنال‌های تنک

²³ Projection onto Convex Sets - POCS

فرض کنید $X \in R^n$ مجموعه ای محدب، بسته، و غیر تهی باشد. برای هر $z \in R^n$ تصویر z روی X را برابر با x ای درون X می‌گیریم که عبارت $\|z - x\|$ را مینیمم نماید، و آن را با $P_X(z)$ نمایش می‌دهیم. به عبارت دیگر تصویر هر بردار z روی X را برابر با نزدیک ترین بردار درون X به z در نظر می‌گیریم. این نوع تصویر کردن ویژگی‌های متعددی دارد که در زیر به بیان برخی از آنها می‌پردازیم.

قضیه (۷-۳): برای هر $z \in R^n$ تصویر یکتایی در X وجود دارد، که آن را با x^* نمایش می‌دهیم.

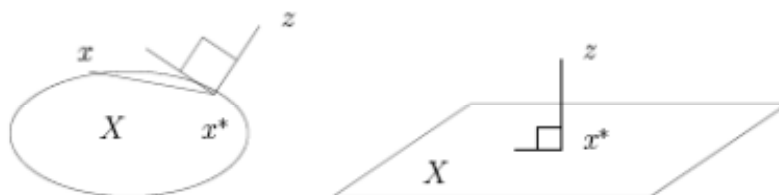
اثبات: تابع $\|z - x\|$ بر حسب x اکیدا محدب و همیشه مثبت است. لذا بنابر قضیه ویراشتراس دارای مینیمم یکتا است. این مینیمم همان x^* است.

قضیه (۸-۳): x^* نقطه یکتایی است که برای آن به ازای هر $x \in X$ داریم $(z - x^*)^T (x - x^*) \leq 0$.

اثبات: از برهان خلف استفاده می‌نماییم. فرض کنید نقطه x وجود دارد به نحوی که برای آن $(z - x^*)^T (x - x^*) > 0$ و تعریف کنید $d = x - x^*$. بنابر تحدب X برای هر $\alpha \in [0, 1]$ داریم $(x^* + \alpha d) \in X$. قضیه کسینوسها نتیجه می‌دهد:

$$\|z - (x^* + \alpha d)\|^2 = \|z - x^*\|^2 + \alpha^2 \|d\|^2 - 2\alpha (z - x^*)^T d$$

واضح است که اگر α به اندازه کافی کوچک انتخاب گردد (به عبارت دقیق تر اگر $\alpha^2 \|d\|^2 - 2\alpha (z - x^*)^T d < 0$) آنگاه داریم $\alpha < 2(z - x^*)^T d / \|d\|^2$. این بدان معنی است که فاصله z تا $x^* + \alpha d$ از فاصله آن تا x^* کمتر است و این با فرض این که x^* تصویر z در X است در تناقض است. به این ترتیب برهان کامل می‌گردد. شکل زیر به منظور نمایش این قضیه به صورت گرافیکی رسم شده است. قضیه می‌گوید که زاویه بین بردارهای $x^* z$ و $x^* x$ هیچ گاه از ۹۰ درجه کمتر نخواهد شد. با توجه به شکل سمت راست واضح است که در حالتی که X یک زیر فضا باشد، این زاویه همواره برابر با ۹۰ درجه است. این مطلب را در قضیه بعد اثبات خواهیم نمود.



شکل ۳-۲- نمایش گرافیکی خاصیت دوم تصویر کردن بر روی مجموعه‌های محدب.

قضیه (۹-۳): اگر X یک زیر فضا باشد، به بیان دیگر اگر ماتریس A وجود داشته باشد به قسمی که $X = \{x \in R^n : Ax = 0\}$ ، آنگاه $(z - x^*) \perp X$.

اثبات: فرض کنید $y \in X$. می‌توان از قضیه قبل دو بردار $x = x^* + y$ و $x = x^* - y$ را به جای x قرار داد و نتیجه گرفت که $z - x^*$ بر هر بردار دلخواه در X عمود است. لذا $(z - x^*) \perp X$.

قضیه (۱۰-۳): عملگر تصویر کردن بر روی مجموعه محدب غیر گسترش یابنده^{۲۴} است. به بیان دیگر

$$\forall x, y \in R^n, \|P_X(x) - P_X(y)\| \leq \|x - y\|$$

اثبات: بنابر قضیه (۸-۳) برای هر $\omega \in X$ داریم:

$$(x - P_X(x))^T (\omega - P_X(x)) \leq 0$$

$$(y - P_X(y))^T (\omega - P_X(y)) \leq 0$$

در معادله اول ω را برابر $P_X(y)$ و در دومی آن را برابر $P_X(x)$ قرار دهید. از تلفیق این دو معادله به دست می‌آید:

$$[(x - y) - (P_X(x) - P_X(y))]^T [P_X(y) - P_X(x)] \leq 0$$

با اندکی جابجایی متغیرها و استفاده از این که برای هر دو بردار a و b داریم $\|ab\| \leq \|a\| \|b\|$ نتیجه می‌شود:

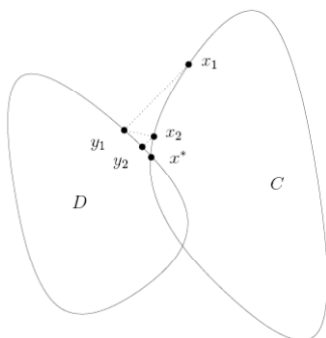
$$\|P_X(x) - P_X(y)\|^2 \leq \|x - y\| \|P_X(x) - P_X(y)\|$$

با تقسیم طرفین این معادله بر $\|P_X(x) - P_X(y)\|$ حکم قضیه اثبات می‌گردد.

کاربرد تصویر کردن بر مجموعه‌های محدب به عنوان یک روش معکوس سازی در نظریه سیستمها به این ترتیب است که فرض کنید بدانیم سیگنال اولیه s که در دست است، عضو یک مجموعه محدب مثل X است. لذا تصویر s روی X همان s است. اگر سیگنال ثانویه را با r نمایش دهیم، طبق قضیه (۷-۳)، فاصله $P_X(r)$ تا s از فاصله r تا s کمتر است. (در مورد سیگنال‌های x و y درون فضای هرمیتی، فاصله به صورت $\int_{-\infty}^{\infty} \|x(t) - y(t)\|^2 dt$ تعریف می‌گردد). لذا می‌توانیم با تصویر کردن سیگنال ثانویه بر این مجموعه محدب به سیگنال جدیدی با فاصله کمتر

²⁴ Non Expansive

از سیگنال اولیه دست یابیم. می‌توان این ایده را تعمیم داد و در صورتی که S عضو چندین مجموعه محدب مجزا باشد از تصویر کردن‌های متوالی سیگنال دریافتی بر روی این مجموعه‌ها جهت ایجاد سیگنالی نزدیکتر استفاده کرد. می‌توان این ایده را به طور نمادین به صورت شکل زیر نمایش داد. واضح است که پس از هر تصویر کردن، سیگنال حاصله به سیگنال اولیه که در ناحیه مشترک دو مجموعه C و D قرار دارد نزدیک تر می‌گردد.



شکل ۳-۳- تصویر کردن‌های متوالی بر روی مجموعه‌های محدب

نسخه‌های دیگری از این روش نیز وجود دارند. یک نمونه از این روش‌های جانبی به این صورت است که سیگنال ثانویه بر مجموعه محدبی که سیگنال اولیه را در بر دارد تصویر گردد، سپس سیگنال حاصله با روشی غیر از روش تصویر کردن باز هم بیشتر به سیگنال اولیه نزدیک گردد، و این فرآیند دو مرحله ای تا رسیدن به سیگنالی بسیار نزدیک به سیگنال اولیه تکرار گردد.

۳-۷- خلاصه

در این فصل روش‌های مبتنی بر محاسبات عددی مورد استفاده در معکوس سازی سیستمها را مورد مطالعه قرار دادیم. می‌توان از این روش‌ها برای تسریع روش‌های مورد استفاده در این گزارش برای تفکیک مودی تعمیم یافته استفاده کرد.

فصل بعد به تحلیل نظری روش معکوس سازی برای تفکیک مودی تعمیم یافته مبتنی بر تقریب‌های متوالی اختصاص یافته است. در این فصل چند شرط لازم و چند شرط کافی برای این کار خواهیم داد.

فصل چهارم

**معکوس سازی با استفاده از تقریب‌های متوالی
مناسب برای تفکیک مودی تعمیم یافته سیگنال‌های
تنک**

۴-۱- مقدمه

در این فصل قصد داریم ثابت کنیم که با استفاده از روشهای تقریب‌های متوالی مبتنی بر Successive Approximations می‌توان معکوس سازی سیستمها را انجام داد، به شرطی که سیستم مربوطه شرایط خاصی (غیر از شرایط ذکر شده در نظریه قالب) را برآورده کند. از قضا شرایط حاکم در مسائل مربوط به تفکیک مودی تعمیم یافته همین شرایط مورد نظر هستند. شرایط لازم و یا کافی برای این که معکوس این سیستم‌ها توسط روشهای مبتنی بر تقریبهای متوالی قابل پیاده سازی باشد را بررسی می‌نماییم. به عبارت دیگر می‌خواهیم توانایی این روشها مطرح شده در ساخت معکوس سیستمهای مربوطه را بررسی کنیم. از این پس هر جا از عبارت "روش تقریب متوالی" یاد می‌شود مقصود روش معکوس سازی با استفاده از تقریب‌های متوالی است. همگرایی این روش هنگامی بسیار زیاد می‌شود که تعداد بسیار زیادی از ضرایب حوزه پردازش برابر با صفر باشد، که متناظر با تنک بودن سیگنال در آن حوزه است.

۴-۲- چند شرط لازم و چند شرط کافی

۴-۲-۱- نمادگذاری‌ها:

نمادگذاری‌هایی که در اینجا استفاده می‌کنیم به قرار زیر هستند.

G : سیستمی که هدف، ساخت معکوس آن است.

$G(x)$: سیگنال اولیه‌ای که ورودی سیستم بوده است، و می‌تواند یک سیگنال گسسته یا پیوسته باشد.

G^{-1} : معکوس G ، به عبارت دیگر $G^{-1}(G(x)) = x$.

y_n ($1 \leq n < \infty$): سیگنال بازسازی شده پس از n بار استفاده از روش تقریب‌های متوالی.

سیستم CI : یک سیستم را CI ^{۲۵} می‌نامیم اگر بتوان معکوس آن را به طور کامل توسط روش تقریب‌های متوالی ساخت.

سیستم CPI ^{۲۶}: سیستمی است که معکوس آن به طور کامل یا به طور جزئی قابل پیاده سازی با روش تقریب‌های متوالی باشد. استفاده از کلمه "جزئی" به این نکته اشاره می‌کند که هر چند روش

^{۲۵} Completely Invertible

^{۲۶} Completely or Partially Invertible

معکوس سازی ممکن است نتواند معکوس سیستم خاصی را به طور دقیق بسازد اما می‌تواند این کار را به طور جزئی انجام دهد، به این معنی که ممکن است با استفاده از هر تعداد تکرار سیگنال ورودی به آن هیچ گاه به صورت کامل باز سازی نشود، اما پس از هر بار تکرار، SNR سیگنال به دست آمده نسبت به SNR سیگنال حاصل شده در مرحله قبلی بالا تر می‌رود. از لحاظ ریاضی

$$SNR(y_n) > SNR(y_{n-1}) \quad 1 \leq n < \infty \quad (۱-۴)$$

هر چند ممکن است با افزایش n به سمت بی نهایت، SNR به بی نهایت میل نکند. واضح است که خانواده سیستم‌های CI ، خود زیر گروهی از سیستم‌های CPI است. در ضمن روشن است که اگر G عضوی از خانواده CI باشد، آنگاه

$$\lim_{n \rightarrow \infty} SNR(y_n) = \infty \quad (۲-۴)$$

تابع (function): G را یک تابع می‌نامیم، اگر یک سیستم علی و بی حافظه باشد.

$awma(y)$:^{۲۷} به معنی y_0 خاصی است که کمترین مقدار قدر مطلق را دارد، به شرطی که متغیر عبارت y روی تمامی عناصر مجموعه S تغییر کند.

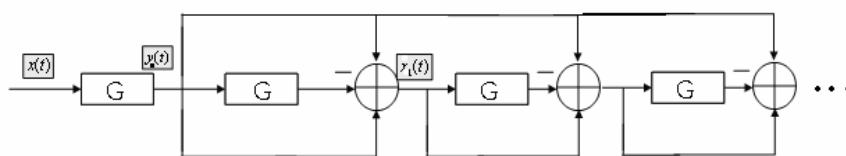
۴-۲-۲- فرمول تقریب های متوالی

اگر قرار دهیم $y_0 = G(x)$ ، آنگاه چنان که در بخش معرفی تئوری تقریب‌های متوالی عنوان شد، به صورتی که در پی می‌آید پیش می‌رویم:

$$y_n = y_{n-1} + \lambda(G(x) - G(y_{n-1})) \quad (۴-)$$

(۳)

که در آن λ پارامتر آرامش است و باید با توجه به مورد تنظیم شود. برای $\lambda = 1$ می‌توانیم روش تقریب های متوالی را همانند شکل (۴-۱) به صورت گرافیکی نشان دهیم.



شکل ۴-۱- نمایش روش تقریب‌های متوالی به صورت گرافیکی

۴-۲-۳- شرط کافی روی G برای این که به خانواده سیستم‌های CPI تعلق داشته باشد

تصور کنید که x تابعی از زمان باشد، و در ضمن به دلایلی که بعداً روشن می‌گردد فرض کنید که پارامتر آرامش را در هر تکرار تغییر می‌دهیم. با توجه به (۴-۳) داریم:

$$y_n(t) - x(t) = y_{n-1}(t) - x(t) + \lambda_n [G(x(t)) - G(y_{n-1}(t))] \quad (4-4)$$

$$e_n^2 = \|y_n(t) - x(t)\|^2 = e_{n-1}^2 + \lambda_n^2 \|G(x(t)) - G(y_{n-1}(t))\|^2 + 2\langle \lambda_n [G(x(t)) - G(y_{n-1}(t))] y_{n-1}(t) - x(t) \rangle \quad (4-5)$$

(۵)

که در آن نمادهای $\langle \rangle$ و $\| \cdot \|$ مطابق زیر تعریف شده اند:

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t) \cdot \bar{y}(t) dt$$

$$\|x(t)\|^2 = \langle x(t), x(t) \rangle$$

علامت bar روی y برای نشان دادن مزدوج مختلط استفاده شده است. تعریف می‌کنیم:

$$A_n = \lambda_n^2 \|G(x(t)) - G(y_{n-1}(t))\|^2 + 2\lambda_n \langle [G(x(t)) - G(y_{n-1}(t))] y_{n-1}(t) - x(t) \rangle \quad (4-6)$$

واضح است که اگر $A_n \leq 0$ ، آنگاه در n امین تکرار SNR افزایش پیدا می‌کند یا حداقل کوچکتر نمی‌شود. یعنی:

$$SNR(y_n) \geq SNR(y_{n-1})$$

حال مجموعه λ هایی را پیدا می‌کنیم که باعث می‌شوند A_n نا مثبت باشد. تعریف می‌کنیم:

$$\alpha = -2 \frac{\langle [G(x(t)) - G(y_{n-1}(t))], y_{n-1}(t) - x(t) \rangle}{\|G(x(t)) - G(y_{n-1}(t))\|^2}$$

با توجه به سهمی بودن تابع نشان داده شده در (۴-۶) نسبتاً به سادگی می‌توان نشان داد که مقادیری از λ که در تکرار n ام می‌توانند استفاده شوند برای این که داشته باشیم

$$SNR(y_n) \geq SNR(y_{n-1})$$

به صورت زیر هستند:

$$\text{If } \alpha > 0 \Rightarrow \text{ the set is } \lambda_n \in (0, \alpha]$$

(۷-۴)

$$\text{If } \alpha < 0 \Rightarrow \text{ the set is: } \lambda_n \in [\alpha, 0)$$

(۸-۴)

از مشتق گیری (۴-۶) نسبت به λ_n و مساوی صفر قرار دادن آن نتیجه می‌شود که بیشترین کاهش A_n نسبت به A_{n-1} به ازای $\lambda_n = \alpha/2$ به دست می‌آید.

علاقه مندیم λ ای را چنان بیابیم که بتوانیم از آن در تمامی تکرارها استفاده کنیم. اگر تعریف کنیم:

$$B = \{r(t) : \|r(t) - x(t)\| \leq \|G(x(t)) - x(t)\|\} \quad -۴)$$

(۹)

آنگاه به شرط برقراری شرط زیر مقصود ما حاصل می‌گردد:

شرط (۱-۴): برای هر $r(t) \in B$ علامت عبارت $\langle [r(t) - x(t)], [G(r(t)) - G(x(t))]\rangle$ یاهمواره مثبت و یا همواره منفی است.

قضیه (۱-۴): با انتخاب λ همانند زیر، با افزایش n همواره SNR بهبود می‌یابد.

$$\lambda = \text{sign}(P(t)) \text{Min}_{r(t) \in B} \frac{|P(t)|}{\|G(r(t)) - G(x(t))\|^2} \quad (۱۰-۴)$$

که در این عبارت $P(t) = \langle [r(t) - x(t)], [G(r(t)) - G(x(t))]\rangle$

اثبات: به وضوح $y_0 = G(x)$ متعلق به B است. فرض کنید که در تکرار n ام هستیم و $y_{n-1}(t) \in B$. با توجه به روش تعریف λ در (۴-۱۰) و با توجه به (۴-۷) و (۴-۸) به دست می‌آوریم:

$$SNR(y_n(t)) > SNR(y_{n-1}(t))$$

(۴-۱۱)

بنابر این اثبات کردیم که اگر شرط (۴-۱) ارضا شود آنگاه نرم خطای خروجی مرحله n ام از نرم خطای مرحله $n-1$ ام کمتر می‌شود. از طرفی B در واقع مجموعه سیگنال‌هایی است که فاصله آنها از x کمتر است از یا مساوی است با فاصله $G(x(t))$ از x . بنا بر فرض ابتدایی ما $y_{n-1}(t) \in B$ و چون اثبات کرده ایم که فاصله y_n از x کمتر است از فاصله y_{n-1} از x ، بنابر این $y_n \in B$ و در ضمن می‌توانیم استقرا را ادامه دهیم. بنابر این ادعای ما به این ترتیب اثبات شد، یعنی به شرط برقراری شرط (۴-۱)، G به خانواده سیستم‌های CPI تعلق دارد.

اگر x یک سیگنال گسسته باشد و G سیستمی که روی سیگنال‌های گسسته عمل می‌کند باشد، پارامترها و نماد گذاری‌ها باید به ترتیب زیر اصلاح گردند:

$$\langle f_1[k], f_2[k] \rangle = \sum_{i=-\infty}^{\infty} f_1[i] \cdot \bar{f}_2[i]$$

$$\|f[k]\|^2 = \langle f[k], f[k] \rangle$$

$$B = \{r[k] : \|r[k] - x[k]\| < \|G(x[k]) - x[k]\|\}$$

شرط (۴-۱) هم به صورت زیر تغییر پیدا می‌کند:

شرط (۴-۱) در حالت گسسته: برای هر $r[k] \in B$ علامت عبارت

$$\langle r[k] - x[k], G(r[k]) - G(x[k]) \rangle$$

به صورت زیر تعریف شود:

$$\lambda = \text{sign}(P[n]) \text{Min}_{r(t) \in B} \frac{|P[n]|}{\|G(r[n]) - G(x[n])\|^2}$$

که در این عبارت $P[n] = \langle r[n] - x[n], G(r[n]) - G(x[n]) \rangle$.

۴-۲-۴- شرط کافی روی توابع برای این که عضو خانواده سیستم‌های CI باشند

فرض کنید G یک تابع است. در این صورت آن را با g نشان می‌دهیم. نکته قابل توجه این است که برای $-\infty < t_0 < \infty$ ، $g(x(t_0))$ تنها به مقدار $x(t)$ در لحظه t_0 بستگی دارد. بنابراین در آن چه در پی می‌آید فرض می‌کنیم که ورودی g یک عدد مثل x_0 است و خروجی آن نیز عدد دیگری است

قضیه (۲-۴): اگر g یک تابع مشتق پذیر و در بازه

$$I = [\min(x_0, g(x_0)), \max(x_0, g(x_0))]$$

یکنوا باشد، آنگاه x_0 از روی $g(x_0)$ قابل بازیابی است، با مقدار

$$\lambda = \text{awma}_{r \in I} \frac{1}{g'(r)} \quad (۱۲-۴)$$

اثبات: اولاً واضح است که هر گاه با یک تابع سروکار داشته باشیم (۳-۴) به صورت زیر در می‌آید:

$$y_n = y_{n-1} + \lambda(g(x_0) - g(y_{n-1})) \quad (۱۳-۴)$$

برای اثبات، ابتدا فرض کنید که g یک تابع اکیدا صعودی تعریف شده در بازه I است، و در ضمن $g(x_0) > x_0$. در این صورت به وضوح $y_1 = g(x_0) \in I$ و $y_1 \geq x_0$. فرض کنید در تکرار n ام هستیم و داریم $y_{n-1} \in I$ و $y_{n-1} \geq x_0$. می‌دانیم که:

$$\exists s \in I : x_0 \leq s \leq y_{n-1} \text{ such that } g(x_0) - g(y_{n-1}) = (x_0 - y_{n-1}) \cdot g'(s)$$

لذا (۱۳-۴) تبدیل می‌شود به:

$$y_n - y_{n-1} = \lambda(x_0 - y_{n-1})g'(s)$$

چون g یک تابع اکیدا صعودی است، $g'(s) > 0$. در ضمن با توجه به (۱۲-۴) داریم $\lambda > 0$ ، بنابراین $\lambda(x_0 - y_{n-1}) \leq 0$ و این نتیجه می‌دهد که

$$y_n < y_{n-1}$$

(۱۴-۴)

با توجه به (۱۲-۴)، $\lambda g'(s) < 1 \Rightarrow y_n - y_{n-1} > x_0 - y_{n-1}$ ، بنابراین داریم:

$$y_n > x_0 \quad (15-4)$$

(۱۴-۴) و (۱۵-۴) به دست می‌دهند $x_0 < y_n < y_{n-1}$ ، بنابراین $y_n \in I$ می‌توان با استقرا کار را دنبال کرد و نشان داد:

$$x_0 < y_n < y_{n-1} \quad 2 \leq n < \infty$$

می‌بینیم که دنباله $\{y_i\}_{i=1}^{\infty}$ یک دنباله اکیدا نزولی است که حد پایینی مثل x_0 دارد. بنابراین این دنباله دارای حد است. این حد را با Z نمایش می‌دهیم. از (۱۳-۴) وقتی $n \rightarrow \infty$ به دست می‌آوریم:

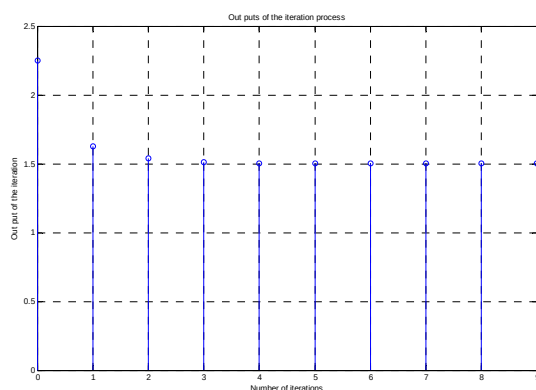
$$z = z + \lambda(g(x_0) - g(z)) \Rightarrow g(x_0) = g(z)$$

اما از آن جا که فرض کرده ایم که g یک تابع اکیدا یکنوا است به دست می‌آوریم $x_0 = z$. این نتیجه می‌دهد:

$$x_0 = \lim_{n \rightarrow \infty} y_n$$

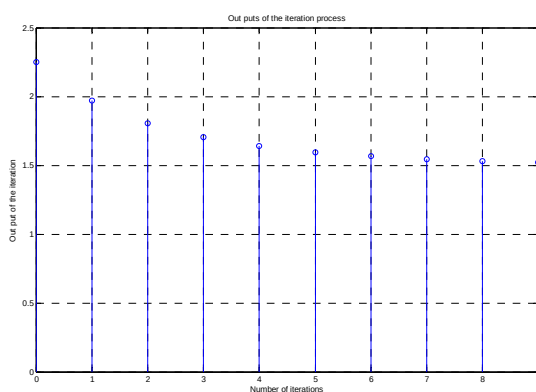
بنابر این قضیه در حالتی که g یک تابع اکیدا صعودی روی I باشد و در ضمن $g(x_0) > x_0$ ثابت شد. خواننده به سادگی می‌تواند این قضیه را به سه حالت باقی مانده نیز تعمیم دهد: یعنی g صعودی باشد و $g(x_0) < x_0$ ، g نزولی باشد و $g(x_0) > x_0$ ، یا g نزولی باشد و $g(x_0) < x_0$.

به عنوان یک مثال، فرض کنید $g(x) = x^2$. این تابع روی بازه $[0, \infty)$ یکنوا است و به سادگی می‌توان دید که شرایط قضیه (۲-۴) را ارضاع می‌کند. فرض کنید ورودی تابع $1/5$ باشد، لذا خروجی آن $2/25$ خواهد بود. برای این تابع، $\frac{1}{g'(r)}$ مساوی با $\frac{1}{2r}$ است و با توجه به قضیه، λ میتواند روی مقدار $\frac{1}{2 \times 2.25}$ یعنی 0.222 تنظیم شود. خروجی روش تقریب‌های متوالی در شکل (۲-۴) رسم شده است.



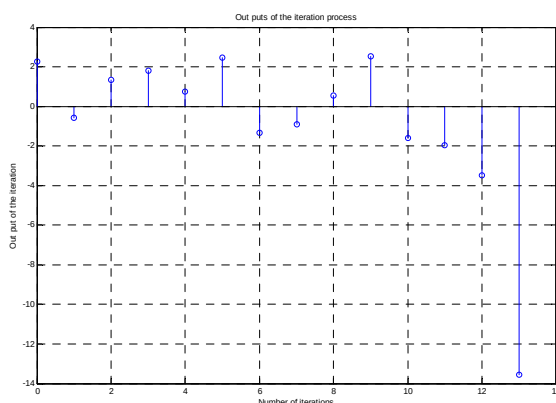
شکل ۴-۲- نتیجه عملیات تقریب‌های متوالی برای $g(x) = x^2$ و $\lambda = 0.222$

برای نشان دادن اثر پارامتر λ ، در نمودار بعدی آن را برابر 0.1 قرار داده ایم. نرخ همگرایی کمتر، به وضوح از شکل مشخص است.



شکل ۴-۳- نتیجه عملیات تقریب‌های متوالی برای $g(x) = x^2$ و $\lambda = 0.1$

اگر λ مقدار بسیار بزرگی انتخاب شود روش تقریب‌های متوالی ممکن است واگرا شود. این حالت در شکل (۴-۴) نشان داده شده است. در این شکل λ برابر با $1/0.1$ انتخاب شده است.



شکل ۴-۴- نتیجه عملیات تقریب‌های متوالی برای $g(x) = x^2$ و $\lambda = 1.01$

حال فرض کنید $x(t)$ که $-\infty < t < \infty$ ، ورودی g باشد. هم چنین فرض کنید a و b وجود دارند چنان که برای هر t داریم $a < x(t) < b$. برای هر α که $a < \alpha < b$ بازه $I(\alpha)$ را به شکل زیر تعریف می‌کنیم:

$$I(\alpha) = [\min(\alpha, g(\alpha)), \max(\alpha, g(\alpha))]$$

قضیه (۴-۳): مجموعه‌ای از شرایط کافی روی g برای این که $x(t)$ قابل به دست آورده شدن توسط روش تقریب‌های متوالی مطرح شده باشد این است که:

اولاً اگر $I_0 = \bigcap_{\alpha=a}^b I(\alpha) \neq \emptyset$ ، آنگاه I_0 ثانیاً g تابعی اکیدا یکنوا و مشتق پذیر روی $\bigcup_{\alpha=a}^b I(\alpha)$ باشد.

در این صورت λ ای که برای انجام کار پیشنهاد می‌شود برابر است با:

$$\lambda = awma \frac{1}{g'(r)}_{r \in I_0}$$

اثبات: اثبات این قضیه تعمیم سر راستی از قضیه (۴-۲) است.

۴-۲-۵- بازسازی x از روی $G(x)$ وقتی G یک ماتریس معکوس پذیر است

فرض کنید که $G(x) = Ax$ ، که در آن A یک ماتریس هرمیتی $N \times N$ و x یک بردار $1 \times N$ است. تعریف می‌کنیم:

$$e_n = x - y_n \quad 1 \leq n \leq \infty$$

به آسانی و با استقرا می‌توان نشان داد که (I) ماتریس واحد $N \times N$ است.

$$e_n = (I - A)(I - \lambda A)^{n-1} x$$

برای این که برای هر بردار x داشته باشیم $\lim_{n \rightarrow \infty} \|e_n\| = 0$ لازم و کافی است که

$$\lim_{n \rightarrow \infty} \|(I - \lambda A)^{n-1}\| = 0$$

فرض کنید که β_i ها ($1 \leq i \leq n$) مقادیر ویژه ماتریس $B = I - \lambda A$ و q_i ها بردارهای ویژه متناظر با آنها باشند. داریم

$$(I - \lambda A)^{n-1} = \sum_{i=1}^N \beta_i^{n-1} q_i q_i^H$$

که در آن H علامت ترانهاد مزدوج^{۲۸} است. لذا بدیهی است که باید داشته باشیم:

$$|\beta_i| < 1 \quad 1 \leq i \leq N$$

(۱۶-۴)

با استفاده از تعریف مقادیر ویژه

$$(I - \lambda A)q_i = \beta_i q_i \quad 1 \leq i \leq N \quad \Rightarrow \quad Aq_i = \frac{1 - \beta_i}{\lambda} q_i$$

بنابراین $\gamma_i = \frac{1 - \beta_i}{\lambda}$ ها ($1 \leq i \leq N$) مقادیر ویژه ماتریس A هستند. داریم $\beta_i = 1 - \lambda \gamma_i$ ، بنابراین (۱۶-۴) به دست می‌دهد که

$$-1 < 1 - \lambda \gamma_i < 1 \quad \Rightarrow \quad 0 < \lambda \gamma_i < 2$$

شرط لازم و کافی برای وجود داشتن λ این است که علامت همه γ_i ها یکسان باشد، که در مورد ماتریس هرمیتی معکوس پذیر این خاصیت برقرار است. در این صورت می‌توانیم قرار دهیم

$$\lambda = \frac{2}{\gamma_{\max}}$$

که $\gamma_{\max}$ مقدار ویژه A با بزرگترین مقدار است.

قضیه (۴-۴): اگر A یک ماتریس معکوس پذیر باشد، x همواره با استفاده از روش معکوس سازی قابل به دست آوردن است.

²⁸ Hermitian

اثبات: چون A معکوس پذیر است هیچ یک از مقادیر ویژه آن برابر صفر نیستند. اگر A هرمیتی باشد ، با توجه به مطالب پیش گفته چیزی برای اثبات باقی نمی ماند. در غیر این صورت می توانیم ابتدا $y'_1 = A^H y_1$ را تشکیل دهیم ، و به وضوح $y'_1 = A^H A x$. می دانیم که $A' = A^H A$ یک ماتریس معین مثبت است و بنابر این تمام مقادیر ویژه آن مثبت هستند. لذا می توانیم سیستم مورد بحث را برابر $G = A'$ قرار دهیم و مطابق با مطالب پیش گفته یک λ مناسب را برای ادامه کار انتخاب کنیم.

۴-۲-۶- شرط کافی روی سیستمهای کانولوتیو برای تعلق به گروه CI

فرض کنید G یک سیستم کانولوتیو باشد، یعنی

$$G(x(t)) = x(t) \otimes h(t) \quad (۱۷-۴)$$

که علامت $\otimes$ برای نشان دادن کانولوشن استفاده شده است. اگر از طرفین (۱۷-۴) تبدیل فوریه بگیریم به دست می آوریم:

$$Y_n(f) = Y_{n-1}(f) + \lambda(H(f)(X(f) - Y_{n-1}(f)))$$

(۱۸-۴)

که در آن از حروف بزرگ برای نشان دادن سیگنال‌های حوزه فرکانس استفاده شده است. با استقرا به سادگی نشان داده می شود که

$$E_n(f) = (1 - H(f))(1 - \lambda H(f))^{n-1} X(f)$$

روشن است که اگر بخواهیم با استفاده از روش تقریب های متوالی $x(t)$ را از روی $G(x(t))$ به دست آوریم، شرط لازم و کافی به این صورت قابل بیان است که باید برای فرکانس هایی که $X(f) \neq 0$ داشته باشیم :

$$|1 - \lambda H(f)| < 1 \quad \Rightarrow \quad \frac{A(f)}{\lambda |H(f)|^2} < 0.5 \quad (۱۹-۴)$$

که در آن $A(f)$ برای نمایش بخش حقیقی $H(f)$ به کار رفته است. ملاحظه می شود که این شرط تنها و تنها وقتی می تواند بر آورده شود که علامت $A(f)$ برای هر f همواره مثبت و یا همواره منفی باشد. در این صورت λ می تواند مطابق زیر انتخاب گردد:

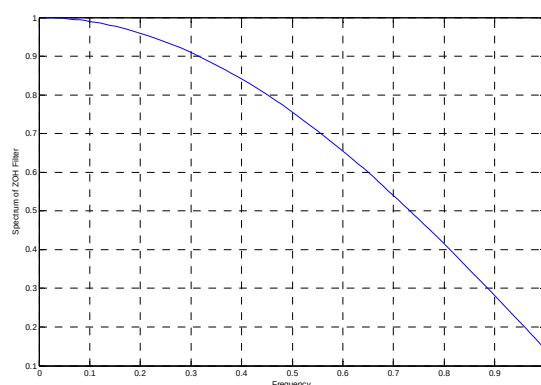
$$\lambda = \underset{f \in R}{\operatorname{awma}} \frac{2A(f)}{\|H(f)\|^2}$$

قضیه (۵-۴): برای هر سیستم کانولوتیو، $x(t)$ می‌تواند با خطایی با انرژی صفر به وسیله روش معکوس سازی به دست آورده شود.

اثبات: اگر (۴-۱۹) برقرار باشد چیزی برای اثبات نمی‌ماند. در غیر این صورت $y_1(t) = y(t) \otimes \bar{h}(t)$ را می‌سازیم که در آن $\bar{h}(t)$ مزدوج مختلط $h(t)$ است. حال تعریف می‌کنیم $h_1(t) = h(t) \otimes \bar{h}(t)$. $H_1(f) = |H(f)|^2$ همیشه (به جز در فرکانس‌هایی که $H(f)=0$) مثبت است. مجموعه فرکانس‌هایی را که برای آنها $X(f) \neq 0$ و $H(f)=0$ است را با D نمایش می‌دهیم. اگر تعداد اعضای D محدود باشد با افزایش تعداد iterationها انرژی سیگنال خطا به صفر می‌گراید زیرا تبدیل فوری سیگنال خطا برای فرکانس‌هایی که عضو D نیستند به صفر می‌گراید و در غیر این صورت به $X(f)$ می‌دانیم که انرژی سیگنال خطایی که تنها در تعداد محدودی از فرکانس‌ها غیر صفر باشد برابر صفر است مگر این که مقادیر غیر صفر آن برابر بی‌نهایت باشند یعنی در طیف آن ضربه داشته باشیم. در این جا فرض می‌شود چنین ضربه‌هایی در طیف سیگنال خطا وجود ندارد و بنا براین قضیه اثبات می‌شود. توجه خواننده را به این موضوع جلب می‌کنیم که بخش‌های (۴-۲-۵) و (۴-۲-۶) روی هم نشان می‌دهند که روش تقریب‌های متوالی ارائه شده همواره توانایی ساختن معکوس سیستم‌های خطی را دارد.

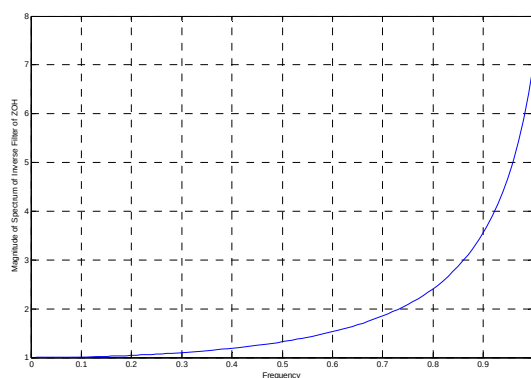
۴-۳- مزایای معکوس سازی سیستم‌ها با استفاده از روش تقریب‌های متوالی

این مطلب که می‌توان با استفاده از سیستم اصلی معکوس آن را ساخت از لحاظ عملی بسیار قابل توجه است. در این جا برای این که خواننده را به اهمیت موضوع واقف گردانیم به بیان یک مثال ساده می‌پردازیم. یک سیستم کانولوشن غیر ایده آل را در نظر بگیرید که از ساختار تقریب پالسی برای نمونه برداری استفاده می‌کند. بنابراین اگر توجه خود را به حوزه فرکانس معطوف کنیم واضح است که اندازه طیف سیگنال اولیه در پنجره ای مطابق روبرو ضرب می‌گردد:



شکل ۴-۵- طیف سیگنال متناظر با پالس حوزه زمان

اولین ایده ای که برای حذف اثر این پنجره فرکانسی به نظر می‌رسد عبور دادن سیگنال در یافتی در گیرنده از فیلتری با اندازه پاسخ فرکانسی به صورت زیر است:



شکل ۴-۶- طیف سیگنالی که اثرات نامطلوب کانولوشن غیر ایده آل را خنثی می‌نماید

واضح است که اندازه این پاسخ فرکانسی در حاشیه آن بزرگ است. اگر نرخ نمونه برداری پایین باشد لازم می‌شود که فیلتری که برای جبران سازی اثر تخریب به کار گرفته می‌شود در حاشیه باند عبور خود مقادیر بسیار بزرگ و در فرکانس‌های پایین مقادیری کوچک را اختیار کند. طراحی چنین فیلتری از لحاظ عملی مشکل خواهد بود. لذا استفاده از روش‌های جایگزین برای حذف اعوجاج ناشی احساس می‌شود. روش تقریب‌های متوالی ارائه شده در این گزارش می‌تواند به عنوان یکی از این جایگزین‌ها به کار رود.

۴-۴- کاربردهای روش تقریب‌های متوالی

روش تقریب‌های متوالی مطرح شده بالقوه کاربردهای بسیار متنوعی دارد. نکته قابل توجه هنگام استفاده از این روش این است که تنها پارامتری که نیاز به تنظیم کردن دارد λ است. بسیاری از سیستم‌ها وجود دارند که رابطه تقریب‌های متوالی مورد بحث نمی‌تواند با داشتن خروجی آنها ورودیشان را به طور کامل به دست دهد، اما می‌تواند سیگنالی بسیار شبیه به ورودی آنها را تولید کند، یا به عبارتی "معکوس تقریبی" آنها را ایجاد نماید. از جمله این سیستم‌ها سیستم‌های تفکیک مودی مورد استفاده در پردازش سیگنال هستند که تقریباً همانی اند. برای نشان دادن این ادعا، پس از بیان اثبات آن که مبتنی بر تئوری تقریب است از شبیه‌سازی استفاده نموده ایم.

۴-۱-۴-۱-۴ اثبات موثر بودن عملیات تقریب‌های متوالی در بهبود کیفیت سیستم‌های تفکیک مودی سیگنال

سیستم‌های تفکیک مودی سیگنال خروجی را به دست می‌دهند که کمابیش شبیه به ورودی آنها است. به عبارت دیگر این سیستم‌ها تقریباً شبیه سیستم همانی عمل می‌کنند. در ضمن معمولاً در طبقه آخر این سیستم‌ها از نوعی فیلتر کردن استفاده می‌شود (و اگر این طور نباشد می‌توان یک فیلتر با پهنای باند عمومی سیگنال‌هایی که معمولاً به وسیله آنها انتقال می‌یابند با آنها سری کرد) و بنابراین سیستم سیستمی هموار است. لذا خروجی این گونه سیستم‌ها می‌تواند (حداقل به طور تقریبی) به صورت یک سری تیلور بر حسب ورودی نوشته شود. به عنوان مثال تصور کنید که قالب $\mathbf{x} = [u_1, u_2, \dots, u_n]^T$ وارد چنین سیستمی شود و در گیرنده قالب $\mathbf{y} = [v_1, v_2, \dots, v_n]^T$ در یافت شود. در این صورت می‌توان نوشت $\mathbf{y} = A_1\mathbf{x} + A_2\mathbf{x}^2 + A_3\mathbf{x}^3 + \dots$ که در آن $\mathbf{x}^i$ بیان گر برداری است که شامل همه حاصل ضرب‌های u_1 تا u_n از درجه i است. به بیان دیگر شامل همه جملاتی به شکل $u_1^{p_1} \times \dots \times u_n^{p_n}$ است که در آن $p_1 + p_2 + \dots + p_n = i$. در ضمن $A_1, A_2, \dots, A_n$ هم بیانگر ماتریس‌های مشتقات جزئی با مرتبه‌های برابر با اندیششان هستند.

این بار نیز برای این که دید دقیق تری در مورد نحوه انتخاب یک عدد ثابت به عنوان پارامتر آرامش برای خواننده ایجاد شود، ابتدا به بحث در مورد انتخاب پارامترهای آرامش متغیر می‌پردازیم. فرض کنید در مرحله m ام از تکرار قرار داریم. بنابر رابطه تقریب‌های متوالی مورد استفاده داریم:

$$\mathbf{y}_{m+1} = \mathbf{y}_m + \lambda_m (G(\mathbf{x}) - G(\mathbf{y}_m))$$

$$\Rightarrow \mathbf{y}_{m+1} - \mathbf{x} = \mathbf{y}_m - \mathbf{x} + \lambda_m (G(\mathbf{x}) - G(\mathbf{y}_m))$$

$$\Rightarrow \mathbf{e}_{m+1} = \mathbf{e}_m + \lambda_m (A_1\mathbf{x} + A_2\mathbf{x}^2 + A_3\mathbf{x}^3 - A_1\mathbf{y}_m - A_2\mathbf{y}_m^2 - A_3\mathbf{y}_m^3 \dots)$$

$$= (I - \lambda_m A_1) \mathbf{e}_m + \lambda_m [A_2 (\mathbf{x}^2 - \mathbf{y}_m^2) + A_3 (\mathbf{x}^3 - \mathbf{y}_m^3) + \dots]$$

$$= (I - \lambda_m A_1) \mathbf{e}_m + \lambda_m [r_1, \dots, r_n]^T$$

که در این روابط از لحاظ نماد گذاری y_n^k تعریفی مشابه x^k دارد.

می‌خواهیم نشان دهیم که تحت شرایطی که در کاربردهای عملی برقرار هستند λ_m می‌تواند چنان انتخاب گردد که نرم l_2 بردار خطا در مرحله $m+1$ ام از نرم متناظر در مرحله m ام کوچکتر شود.

فرض کنید $\mathbf{e}_m = [s_1, \dots, s_n]^T$ و $A_1 = I_{n \times n} + \delta$ که در آن δ ماتریسی $n \times n$ است که درایه‌های آن به اندازه کافی به صفر نزدیک اند (مقصود ما از عبارت "به اندازه کافی" به زودی روشن می‌گردد). داریم

$$\mathbf{e}_{m+1} = [t_1, t_2, \dots, t_n]^T = [(1 - \lambda_m)s_1 + \lambda_m b_1 + \lambda_m r_1, \dots, (1 - \lambda_m)s_n + \lambda_m b_n + \lambda_m r_n]^T$$

که در آن b_i برای هر $1 \leq i \leq n$ برابر حاصل ضرب برداری سطر i ام δ و بردار $\mathbf{e}_m$ است. در واقع چنان که عنوان شد اگر هر یک از سیستم‌های متداول تفکیک مودی را با G نمایش دهیم، G تا حدود زیادی شبیه به سیستم همانی رفتار می‌نماید. در این جا فرض می‌کنیم که میزان این شباهت در رفتار به حدی است که می‌توان فرض کرد برای هر $1 \leq i \leq n$ داریم $|b_i + r_i| \ll |s_i|$. این بدان معنی است که اگر بگیریم $b_i + r_i = \alpha_i s_i$ داریم $|\alpha_i| \ll 1$. برای اثبات این که می‌توان λ_m را چنان انتخاب کرد که نرم l_2 بردار خطا در مرحله $m+1$ ام از نرم متناظر در مرحله m ام کوچکتر شود، کافی است که اثبات شود λ_m می‌تواند چنان انتخاب گردد که برای هر $1 \leq i \leq n$ داشته باشیم $|t_i| < |s_i|$. یا $|1 - \lambda_m + \lambda_m \alpha_i| < 1$ و این معادل اثبات وجود λ_m است چنان که برای هر i داشته باشیم $0 < \lambda_m < \frac{2}{1 - \alpha_i}$. شرط وجود چنین λ_m ای این است که $1 - \alpha_i$ برای تمام i ها مثبت باشد و این با شرط $|\alpha_i| \ll 1$ به وضوح برقرار است.

در عمل، در هر مرحله از تکرار دید خاصی نسبت به این که خطا در آن مرحله چیست نداریم (زیرا اطلاعاتی از سیگنال دقیق اولیه نداریم). از این رو نمی‌توانیم از پارامترهای آرامش متغیر با تکرار استفاده کنیم. از این رو یک پارامتر آرامش به اندازه کافی کوچک (که تعیین آن تا حدودی تجربی است) انتخاب می‌کنیم و در هر مرحله از آن استفاده می‌نماییم. این که λ باید چقدر کوچک انتخاب

شود بستگی به سیستم خاص مورد بحث دارد. شبیه‌سازی‌ها نشان می‌دهند که این یک شگرد عملی موثر است.

۴-۵- جمع بندی

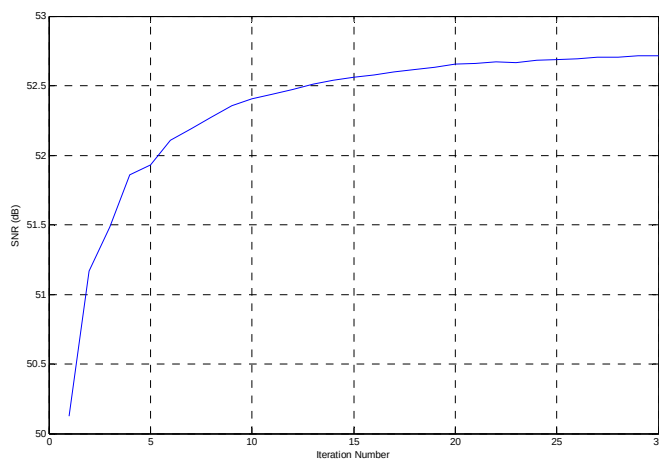
در این فصل مسئله معکوس سازی گروه‌هایی از سیستم‌ها با استفاده از روش تقریب‌های متوالی مبتنی بر تقریب‌های متوالی را به طور نظری بررسی کردیم. در قسمت آخر نشان دادیم که این روش می‌تواند معکوس تقریبی سیستم‌های تفکیک مودی تعمیم یافته و بازسازی (که تقریباً همانی هستند) را بسازد. در این جا باید به طور ضمنی این مطلب برای خواننده روشن شده باشد که چرا معکوس سازی تقریبی یا لااقل خطی سازی سیستم‌های توصیف شده با سری‌های ولترا را به عنوان رگولاتورهای مناسبی برای روش تقریب‌های متوالی خود مطرح کردیم. معکوس سازی تقریبی سیستم توصیف شونده با سری‌های ولترا باعث تبدیل کل سیستم (سیستم اعوجاج ولترایی به همراه معکوس تقریبی یا خطی ساز آن) به یک سیستم تقریباً همانی یا تقریباً خطی می‌گردد و بنابراین سیگنال تقریبی به دست آمده را برای ورود به پروسه تقریب‌های متوالی آماده می‌نماید.

فصل پنجم

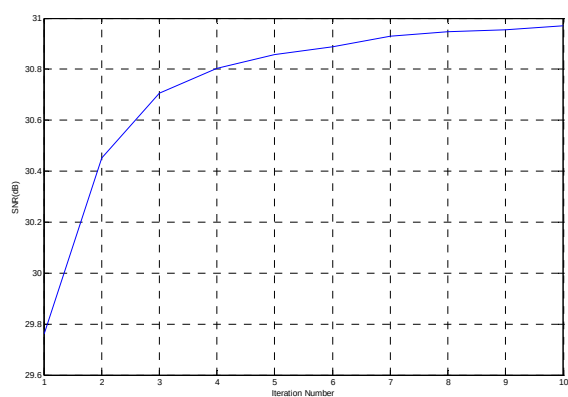
شبیه‌سازی‌ها

در این فصل شبیه‌سازی‌ها ارائه خواهند گردید که در آنها بستر شبیه‌سازی برای سیگنال‌های صوتی آماده گردیده است.

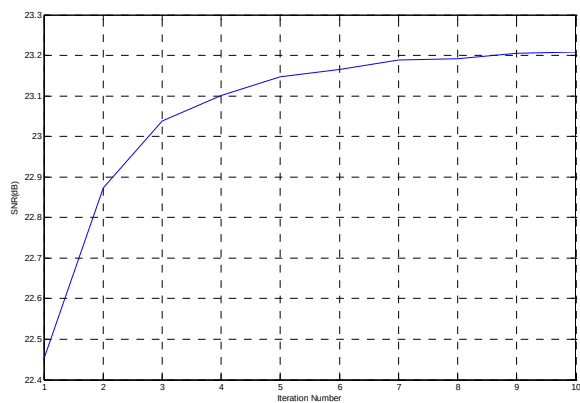
اصواتی را با نرخ ۱۶ کیلو بیت بر ثانیه نمونه برداری کرده ایم و آنها را در دسته ۱ قرار داده ایم. اصوات دیگری از همان فرد را نیز با همین نرخ نمونه برداری کرده ایم و آنها را دسته دوم قرار داده ایم. در شبیه سازی، هر بار یکی از سیگنال‌های گروه یک را وارد کار کرده ایم و ماتریس همبستگی گروه یک را بهنگام کرده ایم. همزمان، یک سیگنال از گروه دوم را هم وارد کار کرده و ماتریس متناظر با گروه دوم را نیز بهنگام نموده ایم. و در گام بعد، بردارهای ویژه تعمیم یافته متناظر با تفکیک مودی تعمیم یافته مورد نظر را با روشهای مندرج در فصول قبل بهنگام کرده ایم. خطای حاصله در هر گام را برابر با نرم یا طول بردار خطای بردار ویژه تعمیم یافته متناظر با بزرگترین مقدار ویژه تعمیم یافته در مقایسه با بردار ویژه تعمیم یافته دقیق تعریف کرده ایم. و نسبت سیگنال به نویز SNR را برابر با ده برابر نسبت طول مقدار بردار مورد بحث به طول خطا بر حسب دسی بل تعریف نموده ایم. با این اوصاف، روند زمانی بهبود این نسبت سیگنال به نویز در صورت استفاده از روشهای مختلف به صورت زیر به دست آمده اند:



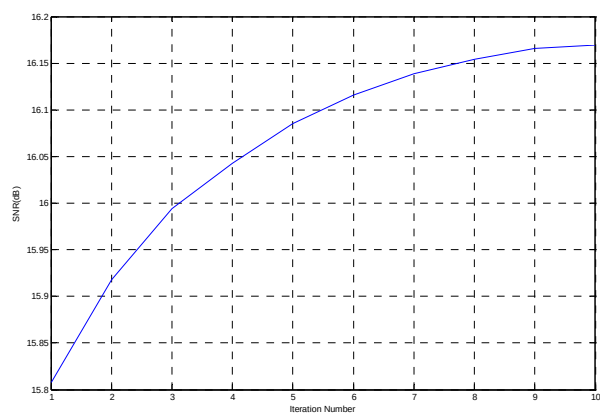
شکل ۵-۱- استفاده از معکوس سازی افقی ماتریس



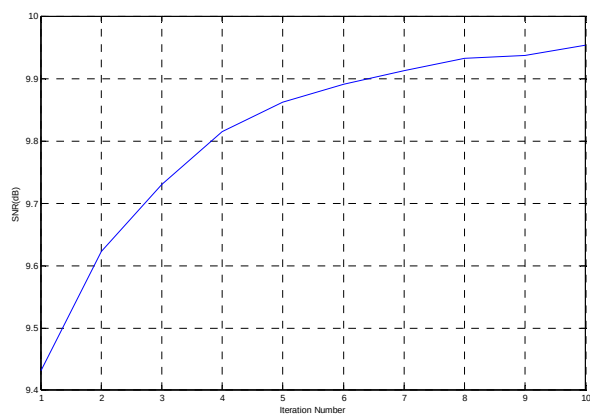
شکل ۵-۲- مبتنی بر استفاده از مدلسازی ولترا



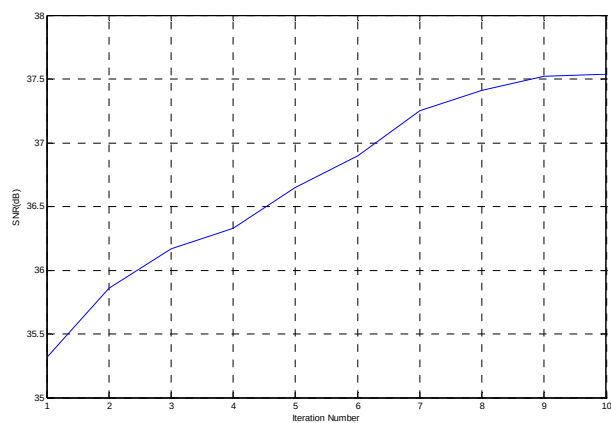
شکل ۵-۳- مبتنی بر فیت کردن در مدل تئوری قالبها



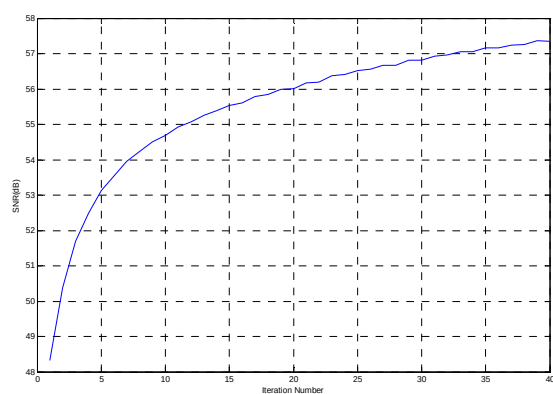
شکل ۵-۴- مبتنی بر روش تسريع چيشف



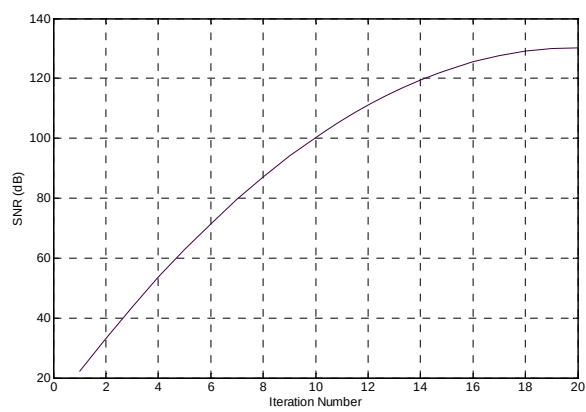
شکل ۵-۵- مبتهی بر روش conjugate gradient



شکل ۵-۶- مبتهی بر روش تقریبهای متوالی با مقدار پارامتر برابر ۰/۲



شکل ۵-۷- مبتهی بر روش تقریبهای متوالی با مقدار پارامتر برابر ۰/۴



شکل ۵-۸- مبنی بر روش تقریبهای متوالی با مقدار پارامتر برابر ۰/۷

منابع

- [1] L. Bradbury and J. W. L. Wan, "A spectral k-means approach to bright field cell image segmentation", "n proc. IEEE Int. Conf. EMBS Soc., Sept. 2010, pp. 4748-4751.
- [2] F. Cerda, S. Chen, J. Bielak, J. H. Garrett, P. Rizzo, and J. Kovacevic, "Indirect SHM by means of sparse representation", "n Proc. Structures Congr., Pittsburgh, PA, May 2013.
- [3] Ray Maleh and Anna Gilbert, "Multichannel image estimation via simultaneous orthogonal matching pursuit", IEEE Workshop on Statistical Signal Processing (SSP), Madison, Wisconsin, August 2007.
- [4] Jianwei Ma and Francois-Xavier Le Dimet, "Deblurring from highly incomplete measurements for remote sensing", IEEE Trans. Geoscience and Remote Sensing, 47 (3), 792-802, 2009.
- [5] Xiao Xiang Zhu and Richard Bamler, "Within The Resolution Cell: Super-Resolution in Tomographic SAR Imaging", Proceedings of the IEEE International Geoscience & Remote Sensing Symposium (IGARSS), Vancouver, Canada, 2011.
- [6] A. Dremeau, C. Herzet and L. Daudet, "Boltzmann machine and mean-field approximation for structured sparse decompositions", IEEE Trans. on Signal Processing, vol 60, pp. 3425-3438, July 2012.
- [7] V. M. Patel, R. Maleh, A. C. Gilbert, and R. Chellappa, "Gradient-based image recovery methods from incomplete Fourier measurements", IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 94-105, Jan. 2012.
- [8] Gongguo Tang and Arye Nehorai, "Lower bounds on mean-squared error for low-rank matrix reconstruction" IEEE Trans. Signal Processing, vol. 59, no. 10, pp. 4559-4571, Oct. 2011.
- [9] Joshua Trzasko, Armando Manduca, and Eric Borisch, "Highly undersampled magnetic resonance image reconstruction via homotopic L0-minimization", IEEE Trans. Medical Imaging 28(1): 106-121, 2009.
- [10] Alfred poltman, "Linear Algebra", wiley, 1991.
- [11] g. bibes, r. ouvrard, p. coirault, "a regularization method for nonlinear inverse problems by using a volterra model", European Control Conference, 2005.

Abstract:

In this research, after a review of the literature and existing algorithms for generalized eigen-decomposition, we will improve some of them for deployment in the case of sparsely representable signals. We will also design some new algorithms for this purpose, and will analyze them. We will then study the behavior of these algorithms when applied to real applications, and show that their output in practice is also considerable.

Keywords:

Generalized eigen-decomposition, signal, principal modes



Iranian Research Institute for Information Science and Technology

Research Project

**Design of Adaptive Algorithms for
Generalized Eigen-Decomposition of
Signals Having Sparse Representations**

Mohammad Mahdi Jahani Yekta

December 2013