بکارگیری برنامه‌ریزی درجه دوم متوالی و نُرم صفر هموارشده

برای بازیابی سیگنال‌های تُنک نویزدار

**محمّدسعید علمداری1\*، مسعود فاطمی2، ابوذر غفاری3**

1. مربی، دانشکده ریاضی کاربردی، دانشگاه خواجه نصیر الدین طوسی، تهران
2. دانشیار، دانشکده ریاضی کاربردی، دانشگاه خواجه نصیر الدین طوسی، تهران
3. استادیار، دانشکده مهندسی برق، دانشگاه علم و صنعت ايران، تهران، ایران

چکیده :

نمایش تُنُک کاربردهاي زیادی در پردازش سیگنال و تصویر دارد که از آن جمله می‌توان به کاربرد در بازسازی تصاویر پزشکی، تقویت و فشرده‌سازی تصاویر، جداسازی سیگنال، پردازش سیگنال آرایه و رادار اشاره کرد. این مهم سبب شده تا محققان از انواع روش‌هاي نمایش تُنُک بهرمند و در جهت حل مسائل بهینه‌سازی از نرم‌های مختلف استفاده نمایند. در این مقاله ابتدا روش‌های مختلف حل نمایش تُنُک بررسی می‌شوند و در ادامه روشی جدید و کارا برای بازیابی سیگنال‌های تُنک نویزدار با بهرمندی از برنامه‌ریزی درجه دوم متوالی و نُرم صفر هموارشده ارائه می‌شود. نتایج آزمایشات انجام شده، نرخ موفقیت بالای روش پیشنهادی در مقایسه با سایر روش‌های بهینه‌سازی نمایش تُنُک را نشان می‌دهد.

واژه های کليدي: بهینه‌سازی، برنامه‌ریزی درجه دوم متوالی، نرم صفر هموارشده، نمایش تُنُک سيگنال.

1. مقدمه

در بهینه‌سازی تُنُک هدف آن است که از بین تعداد زیادي سیگنال پایه که در حالت کلی تعدادشان خیلی بیشتر از بعدشان است، کمترین تعداد براي نمایش یک سیگنال انتخاب شود. تبدیل‌های خطی در ریاضیات و علوم مهندسی جایگاه ویژه‌ای دارند و می‌توان آن‌ها را در فضا با بعد متناهی به کمک دستگاه معادلات خطی نمایش داد. تبدیل‌های خطی کامل در پردازش سیگنال و تصویر به خصوص در زمینه‌های فشرده‌سازی و کاهش نویز سیگنال با اهداف دستیابی به نمایش هر چه تُنُک تر سیگنال، کاربرد دارند.

در سال‌های اخیر، استفاده از نمایش تنک سیگنال در کاربردهایی مانند نمونه برداری، فشرده‌سازی و نویززدایی مورد توجه بسیار قرار گرفته است. موفقیت نمایش تنک در این کاربردها، از آنجا ناشی می‌شود که اکثر سیگنال‌های طبیعی مانند تصویر و گفتار دارای نمایشی تنک هستند. نمایش سیگنال‌ها در پایه مناسب، همواره به عنوان یک گام اساسی در شناخت ویژگی‌ها و تفسیر اطلاعات از سیگنال‌ها مورد توجه بوده است. منظور از نمایش تُنُک، نمایشی از سیگنال است که در آن قسمت عمده‌ای از اطلاعات اساسی سیگنال در تعداد کمی از ضرایب متمرکز می‌شوند. گرچه تعریف و معیاری برای تُنُک بودن وجود ندارد، اما دریک بیان ساده تُنُکی به معنای صفر بودن بیشتر مولفه‌های سیگنال است و در بیان عملی یعنی اکثر مولفه‌های سیگنال کوچک بوده و تنها چند مولفه سیگنال دارای مقادیر قابل توجه‌ای باشند.

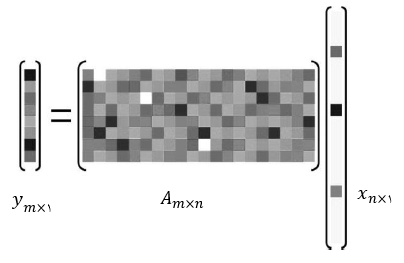
سیگنال‌ها با بردارهای 1× m نشان داده می‌شوند، یعنی فضای سیگنال همان است. یک سیگنال مانند y را در نظر بگیرید، اگر بردارهای یکه فضای را با نشان دهیم، آنگاه سیگنال بصورت زیر نشان داده می‌شود:

|  |  |
| --- | --- |
|  |  |

که در آن ها مولفه‌های بردار y هستند. عبارت فوق، در واقع نمایش بردار y برحسب خانواده بردارهای است. در اصطلاح یک خانواده از سیگنال‌های معلوم را دیکشنری و سیگنال‌های عضو آن را اتم می‌نامند.

دیکشنری ساده‌ترین دیکشنری حوزه زمان است که از اتم‌های کاملا محلی در زمان یعنی سیگنال ضربه یا دیراک و انتقال یافته‌های آن تشکیل شده است. لذا هر سیگنال در ساده‌ترین شکل خود، یعنی هنگامی که آن را بصورت برداری مانند y با مولفه‌های نشان می‌دهیم نمایشی بر حسب یک دیکشنری ساده (یعنی دیکشنری دیراک) است که در آن مولفه‌هایضرایب این نمایش یا بسط می‌باشند.

اکثر سیگنال‌های طبیعي در یک یا چند حوزه مانند زمان، فرکانس، موجک و ... دارای نمایش تُنُک هستند، به این معنا که مي‌توان اطلاعات موجود در آن‌ها را با استفاده از تعداد اندکي ضریب در یک حوزه نمایش خاص بیان کرد. مثلاً اگرچه تصاویر در حوزه مكاني بسیار انبوه یا غني و با اطلاعات زیاد به نظر مي‌رسند، اما در حوزه فرکانس دارای اطلاعات فشرده و به اصطلاح تُنُک هستند، به طوریکه اغلب ضرایب فرکانسي آن‌ها صفر یا نزدیک به صفر است. همین موضوع در مورد بسیاری از انواع دیگر سیگنال‌ها هم صادق است، به طوریکه مي‌توان سیگنالي بسیار حجیم به طول N را تنها با k ضریب که k<Nاست نمایش داد. به عنوان مثال، در شكل1 نمایش تُنُک سیگنال y به وسیله ماتریس حسگر A نشان داده شده است.



شکل 1. نمایش تُنُک سیگنال y به وسیله ماتریس حسگر A .

دراین شکل سیگنال y با ابعاد 1× mبه وسیله ماتریس حسگر A با ابعاد n× m (m<n)نمونه برداری مي‌شود و سیگنال x که در نمایش تنک خیلی کوچکتر از y می‌باشد را نتیجه مي‌دهد. سیگنال اندازه گیری شده x را مي‌توان به گیرنده ارسال کرد و اگر به میزان کافي تُنُک باشد و ماتریس حسگر A شرایط خاصي را ارضا کند، مي‌توانy را از روی x بازسازی کرد. ماتریس A درواقع دربردارنده تمام پایه‌هایي است که برای نمایش داده‌ها استفاده مي‌شود. این ماتریس دیكشنری نامیده مي‌شود و تمام فضای برداری مربوط به داده‌ها را پوشش مي‌دهد. هرکدام از ستون‌های ماتریس دیكشنری را یک اتم مي‌نامند. درصورتي که تعداد اتم‌های دیكشنری برابر بعد فضای برداری باشد آنگاه آن دیكشنری را یک دیكشنری کامل و در این صورت هرکدام از داده‌ها نمایشي یكتا با استفاده از اتم‌های دیكشنری خواهد داشت. اگر تعداد اتم‌های یک دیكشنری کامل را بیشتر کنیم، به آن دیكشنری فراکامل گفته مي‌شود که نمایش سیگنال با آن دارای جواب یكتا نخواهد بود. برای درک بهتر این موضوع دستگاه معادله خطي (1) را در نظر بگیرید.

|  |  |
| --- | --- |
| (1) |  |

به دلیل اینكه ماتریس دیكشنری فراکامل است دستگاه معادلة خطي در پیدا کردن y نامعین است و دارای بي‌شمار جواب خواهد بود. به طور کلی الگوریتم‌های که برای یافتن x با مشاهده y و A بکار مي‌روند، الگوریتم‌های بازیابي تُنُک گفته مي‌شود که در بخش بعدی بطور کامل شرح داده می‌شوند. از آنجایي که به دنبال تُنُک ترین پاسخ برای معادله (1) هستیم، مي‌توانیم مسئله را به فرم یک مسئله بهینه‌سازی به فرم (2) بیان کنیم.

|  |  |
| --- | --- |
| (2) |  |

که در آن معرف تعداد درایه‌های غیرصفر بردار x است و به طورکلي مسئله بالا به معني یافتن بردار x است که دارای کمترین مؤلفه غیرصفر باشد و یا به عبارت دیگر تُنُک ترین x ممكن بدست آید.

الگوریتم‌های بازیابي تُنُک مبتنی بر نرم صفر هموارشده در مقایسه با روش‌های مبتنی بر حداقل‌کردن نرم صفر یا نرم یک، علاوه بر سرعت بیشتر دارای ویژگی‌های مطلوبی از قبیل پیوستگی، مشتق پذیری، سادگی و رفتار عددی مناسب می‌باشند که در بخش دوم شرح داده می‌شود. در این مقاله بر اساس نرم صفر هموارشده روش کارآمدی برای حل (2) پیشنهاد می‌شود که ویژگی‌های آن به شرح زیر است:

1. خواص همگرایی سراسری مطلوبی دارد.
2. از رویکرد جدیدی برای حل زیرمسایل درجه دوم بهره می‌برد.
3. کاهش پیچیدگی محاسباتی مساله را در پی دارد.
4. مساله را در حضور نویز حل می‌کند که بر اساس آن نتایج عددی مطلوبی را به همراه دارد.

ساختار کلی مقاله به شرح زیر است: در بخش2 چهار الگوریتم پرکاربرد در حل نمایش تُنُک بررسی می‌شود. روش برنامه‌ریزی درجه دوم متوالی و روش پیشنهادی در بخش3 آورده شده است. بخش4 نتایج عددی مقاله و مقایسه روش پیشنهادی با سایر الگوریتم‌های شناخته شده در زمینه بازیابی سیگنال‌های تُنک را در قالب سه آزمایش نشان می‌دهد و در نهایت نتیجه گیری در بخش5 آورده شده است.

2. الگوریتم‌های حل نمایش تنک

چالش اصلی در یافتن ضرایب تنک، حل دستگاه معادلات خطی است که برای بازسازی نمایش تنک جهت حل این دستگاه‌ها روش‌های مختلفی وجود دارد. گروهی از آن‌ها، روش‌های حریصانه نام دارند که در آن، بصورت قدم به قدم یک یا چند اتم که بیشترین همبستگی با باقیمانده مربوط به نمایش سیگنال دارند انتخاب و با استفاده از آن‌ها، این باقیمانده را بروز می‌کنند. اساس این روش‌ها مبتنی بر تخمین مرحله به مرحله سیگنال با استفاده از اتم‌های دیکشنری است. به طورکلی این روش‌ها نسبت به سایر الگوریتم‌های نمایش تنک، سریع تر می‌باشند. گروه دیگری از روش‌های بازسازی نمایش تنک، روش‌های رهاسازی مبتنی بر بهینه‌سازی نامیده می‌شوند.

هدف استفاده از روش‌های بازسازی نمایش تنک و یافتن تُنُک ترین پاسخ ممکن برای دستگاه معادلات خطی (1) می‌باشد، لذا مي‌توان مسئله به فرم مسئله بهینه‌سازی (2) در نظر گرفته شود. اگر در معادلة (2) میزان تُنُک بودن را به صورت یک قید یا محدودیت به مسئله اضافه کنیم آنگاه مسئله به فرم (3) درمي‌آید:

|  |  |
| --- | --- |
| (3) |  |

درواقع در اینجا هدف از حل مسئله، کم کردن خطای حاصل از سیگنال بازسازی شده با نمایش تُنُک و سیگنال اولیه است. به خاطر اینكه حل مسئله در حالت کلي غیرممكن است ناگزیر باید مسئله را به فرم معادله (3) حل کنیم و چون خطا و نویز بازسازی اجتناب ناپذیر است در این حالت بردار x به صورت تخمیني بیان شده و معادله اولیه (1) به فرم معادله (4) درمي‌آید.

|  |  |
| --- | --- |
| (4) |  |

در این معادله نویز نمایش نامیده مي‌شود و دارای قید است. با حضور نویز پاسخ‌های تُنُک مسئله (4) را مي‌توان به صورت تقریبي با حل مسئله بهینه‌سازی (5) به دست آورد.

|  |  |
| --- | --- |
| (5) |  |

هر چند روش حل مسائل بهینه‌سازی تعریف شده در بالا با استفاده از نُرم صفر مي‌تواند از روش‌های یافتن پاسخ تُنُک با استفاده از ماتریس به دست آید ولي همچنان پاسخ مسئله غیرقطعي و به طورکلي مستلزم حل یک مسئله NP-HARD است. در ]1[ نشان داده شده است که پاسخ مسئله نُرم صفر را مي‌توان با استفاده از نُرم یک نیز به دست آورد که به مسئله بهینه‌سازی آن گفته مي‌شود. در بسیاری از مسائل کاربردی پاسخ مسائل و معادل هستند و مسئله مي‌تواند تقریب قابل قبولي برای یافتن نمایش تُنُک سیگنال باشد که بصورت (6) تعریف می‌گردد.

|  |  |
| --- | --- |
| (6) |  |

اغلب ساختارهای مشهور نمایش تُنُک بر مبنای نُرم یک بر اساس مسئله بهینه‌سازی (7) حل مي‌شوند:

|  |  |
| --- | --- |
| (7) |  |

یكي دیگر از نرم‌هایي که توسط آن پاسخ تُنُک را به دست مي‌آورند نرم دو است. پاسخ‌های به دست آمده توسط این نرم دارای دقت زیادی نبوده و ممكن است که به اندازه کافي تُنُک نباشند.

در این بخش چهار الگوریتم پرکاربرد در حل نمایش تُنُک به صورت مختصر بیان می‌شوند:

1.2. الگوریتم‌ جست‌وجوی‌ پایه‌

یکی‌از معروف‌ترین‌ ابزارها براي یـافتن‌ پاسـخ‌ تُنُک‌، الگـوریتم‌ جست‌وجوي پایه [[1]](#footnote-1)‎است ‌که‌ در آن‌ به‌ جـاي نـرم‌ صـفر، از حداقل‌ کردن‌ نرم‌ یک‌، مطابق‌ معادله‌ (6)، استفاده‌ می‌شود ]2[. مسئله بهینه‌سازي ‌(6) به‌ راحتی‌ قابل‌ تبدیل‌ به‌ یک‌ مسئله خطی‌ است‌. به‌ همین‌ دلیل‌ با روش‌هاي مختلفی‌ که‌ بـراي مسائل خطی‌ پیشنهاد شده‌ قابل‌ حل‌ است‌. ثابـت‌ شـده‌ کـه‌ اگـر تعداد عناصـر مخـالف‌ صـفر در پاسـخ‌ تُنُک‌ از حـدي کـه‌ بـه‌ همبستگی‌ میان‌ اتم‌ها (ستون‌هاي) دیکشنري A وابسته‌ است‌ کم‌تر باشد، پاسخ‌ رابطه‌ (6) دقیقا برابر با پاسخ‌ تُنُک‌ بـا حـداقل‌سازي نرم‌ صفر است‌.

2.2. الگوریتم‌ جستجوي تطابقی‌

یکی‌از مهم‌ترین‌ ایرادهاي الگوریتم‌ جست‌وجوی‌ پایه‌ زمان‌ مورد نیاز آن‌ اسـت‌ که‌ معمولا با زیاد شدن‌ ابعاد مسئله مشـکل‌ساز می‌شـود. یکی‌از سریع‌ترین‌ روش‌ها براي تخمـین‌ پاسـخ‌ تُنُک‌، الگـوریتم‌ جست‌وجوي تطابق ‌[[2]](#footnote-2)‎ نام‌ دارد ]3[. این روش ابتدا توسط ژانگ و ملت مطرح شد. ذات‌ این‌ روش‌ حریصانه‌ است‌ و تلاش‌ می‌کنـد تـا در هـر گام‌ تنها ضریب‌ یکی‌از اتم‌ها را مشخص‌ کنـد. منظـور از اتم‌هـا ستون‌هاي ماتریس‌ دیکشنري است‌. به‌ عبارت‌ دیگر، در گام‌ اول‌ اتمی‌ که‌ بیش‌تـرین‌ مشـابهت‌ بـا ســیگنال‌ آزمــون‌ را دارد، انتخــاب‌ و ضــریب‌ آن‌ محاسبه می‌شود. در گام‌ بعد، باقی‌مانده‌ سیگنال‌ آزمون‌ و اتم‌ اول‌ با بقیـه‌ اتم‌ها مقایسه‌ و مجددا مشابه‌ترین‌ اتم‌ انتخاب‌ می‌شـود. بـه‌ همین‌ ترتیب‌ در هر گام‌ ضریب‌ یکی‌از اتم‌ها تعیـین‌ می‌شـود تـا جایی‌ که‌ خطاي بازنمایی‌ سیگنال‌ آزمون‌ از حدي کم‌تر شـود یـا آنکه‌ تعداد مشخصی‌ از اتم‌ها داراي ضریب‌ مخالف‌ صفر شوند. این الگوریتم‌ به‌ سبب‌ آنکه‌ در هر گام‌ به‌ یک‌ جست‌وجـوي ساده‌ نیازدارد، معمولا بسیار سریع‌ است‌. اما بـه‌ دلیـل‌ حـریص‌ بودن‌ تضمینی‌ وجود ندارد که‌ پاسخ‌ نهایی‌ مشابه‌ با پاسـخ‌ تُنُک‌ باشد و به جواب بهینه همگرا شود.

3.2. الگوریتم‌ جستجوي تطابقی‌ متعامد‌

به‌ منظور بهبود الگوریتم جستجوي تطابقی ‌ با حفـظ‌ سرعت‌ بالاي آن‌، الگوریتم‌هاي متنوعی‌ پیشـنهاد شـده‌ کـه‌ معروف‌ترین‌ آن‌ الگوریتم جستجوي تطابقی متعامد[[3]](#footnote-3)‎ نام‌ دارد ]4[. در روش جستجوي تطابقی متعامد در هر مرحله ضرایب ستون هاي فعال از ماتریس دیكشنری به صورت مستقل از نتایج مراحل قبل انتخاب می‌شوند و از نتایج قبلی تنها در یافتن مکان مؤلفه هاي غیرصفر استفاده می‌شود. یعنی در هر مرحله بعد از مشخص شدن اتم جدید، از تمام ضرایب قبلی صرف نظر شده و ترکیبی خطی از تمام اتم‌هاي انتخاب شده تا این مرحله، محاسبه می‌شود که کمترین خطا را براي بازسازي سیگنال آزمون داشته باشد. در واقع در این الگوریتم بر خلاف الگوریتم جستجوي تطابقی براي بروز کردن نمایش تُنُک، سیگنال روي زیر فضاي تولید شده توسط اتم‌هاي انتخاب شده تا آن مرحله، تصویر می‌شود.

4.2. الگوریتم‌ نرم‌ صفر هموارشده‌‌

یکی‌ دیگر از الگوریتم‌هاي مناسب‌ براي یافتن‌ پاسخ‌ تُنُک‌، نـرم‌ صفر هموارشده‌[[4]](#footnote-4)‎ نام‌ دارد کـه‌ توسـط‌ بابـایی‌زاده‌ معرفـی‌ شده‌است ]5 .[در این‌ روش‌ تلاش‌ بـر آن‌ اسـت‌ تـا از حـداقل‌ کردن‌ خود نرم‌ صفر استفاده‌ شود. ایرادي که‌ براي حداقل‌ کـردن‌ نرم‌ صفر وجود دارد آن‌ است‌ که‌ تابع‌ نرم‌ صفر پیوسته‌ نیست‌. به‌ همین‌ دلیل‌، پیشنهاد شده‌ تا به‌ جاي نرم‌ صـفر از تـابع‌ همـواري استفاده‌ شود که‌ تقریبی‌ از نرم‌ صـفر را نتیجـه‌ می‌دهـد. نرم‌ صـفر بـردار تعـداد مؤلفـه‌هاي غیرصفر این‌ بردار است‌. به‌ عبارت‌ دیگر اگر فرض‌ کنیم‌:

|  |
| --- |
|  |

آنگاه‌ نرم‌ صفر برابر با است‌، لذا ناپیوستگی‌ نرم‌ صفر معلول‌ ناپیوسـتگی‌ تـابع‌ اسـت‌. اگر این‌ تابع‌ با تابعی‌ همـوار جـايگـزین‌ شـود، آن‌گـاه‌ تخمـین‌ هموار نرم‌ صفر به‌ دسـت‌ می‌آیـد. توابـع‌ گونـاگونی‌ بـراي ایـن‌ منظور می‌توان‌ استفاده‌ نمود کـه‌ مناسـب‌ترین‌ انتخاب‌، قرینه‌ تابع‌ گوسی‌ با میانگین‌ صفر است‌. در این‌ صورت‌ داریم‌:

|  |
| --- |
|  |

به وضوح مشخص است که اگر خیلی کوچک باشد، آنگاه‌ :

درنتیجــه‌ ‌. بنــابراین‌، اگر فرض‌ کنیم‌ داریم:

|  |  |
| --- | --- |
|  |  |

آنگاه‌ می‌تواند تخمین‌ نرم‌ صفر باشد، یعنی:

|  |  |
| --- | --- |
| (8) |  |

از آنجا که‌ این‌ تابع‌ جدید هموار است‌ به‌ سادگی‌ می‌توان‌ آن‌ را بهینه‌سازي نمود. اگر را در رابطه‌ (8) به‌ سمت‌ صفر میـل‌ دهیم‌، تقریب‌ به‌ تساوي تبدیل‌ می‌شود. بنـابراین‌ بیشـینه‌ کردن‌ تابع‌ براي کوچک‌ معادل‌ بـا حـداقل‌ کـردن‌ نـرم‌ صفر است‌. نکته مهم در استفاده از نُرم صفر هموارشده، پارامتر هموارسازی است. تابع به ازای مقادیر کوچک تخمین مناسب تری ارائه می‌کند ولی به شدت ناهموار و دارای تعداد زیادی بیشینه محلی خواهد بود. از طرفی به ازای مقادیر بزرگتر تابع هموارتر و دارای تعداد کمتری بیشینه محلی خواهد بود، لذا برای جلوگیری از به تله افتادن الگوریتم در بیشینه محلی، دنباله کاهشی از را انتخاب و در هر مرحله بیشینه تابع را می‌یابند.

3. مروری بر روش برنامه‌ریزی درجه دوم متوالی و روش پیشنهادی

روش برنامه‌ریزی درجه دوم متوالی[[5]](#footnote-5)‎ یک روش شناخته شده و کارا برای حل مسائل بهینه‌سازی غیرخطی، غیرمحدب و هموار است. این روش اولین بار توسط ویلسون در سال 1963 پیشنهاد و در ادامه توسط هان و پاول رایج شد. رویکرد این روش بدین صورت است که در هر تکرار، برای محاسبه جهت جدید d مسئله بهینه‌سازی درجه دوم حل می‌گردد و در آن از تقریب درجه دوم تابع هدف و تقریب خطی درجه اول قیود مسئله اصلی استفاده می‌شود. سرعت همگرایی زبرخطی و نتایج همگرایی قوی آن، محققان زیادی را به تحقیقات گسترده در این زمینه سوق داده است ]8-6 [. مسئله بهینه‌سازی مقید (9) را در نظر بگیرید:

|  |  |
| --- | --- |
| (9) |  |

برای محاسبه جهت جدید زیر مسئله ناحیه اعتماد درجه دوم (10) در نظر گرفته می‌شود:

|  |  |
| --- | --- |
| (10) |  |

که در آن شعاع ناحیه اعتماد، گرادیان تابع هدف و هسین تابع می‌باشد. بصورت دقیق تر داریم:

|  |  |
| --- | --- |
| (11) |  |

|  |  |
| --- | --- |
| (12) |  |

که در آن ماتریس قطری با عناصر می‌باشد. نمادهای به عنوان کاهش پیش بینی شده و به عنوان کاهش واقعی بصورت زیر در نظر گرفته می‌شوند:

|  |  |
| --- | --- |
| (13) |  |

|  |  |
| --- | --- |
| (14) |  |

رابطه (15)، شرط کاهش کافی در تابع هدف نامیده می‌شود که در آن است.

|  |  |
| --- | --- |
| (15) |  |

در این مقاله روشی جدید و کارا برای بازیابی سیگنال‌های تُنک نویزدار پیشنهاد می‌گردد. در روش پیشنهادی برای حل مسئله بهینه‌سازی، بجای نرم صفر از نـرم‌ صفر هموارشده‌ استفاده می‌کنیم و بجای حل مسئله ناهموار (2) تقریب همواری از آن به فرم (9) لحاظ می‌شود. زمانی که مسئله در معرض نویز باشد، بجای مسئله (9) ، مسئله (16) در نظر گرفته می‌شود.

|  |  |
| --- | --- |
| (16) |  |

ابتکار روش پیشنهادی در این است که مسئله (16) به یک مسئله بهینه‌سازی با قیود برابری تبدیل می‌شود و لذا قید نابرابری حذف و شرط جدیدبه تابع هدف اضافه می‌گردد که در آن پارامتر هموارسازی می‌باشد. بنابراین مسئله به فرم (17) تبدیل می‌شود:

|  |  |
| --- | --- |
| (17) |  |

در اینجا، تخمین نویز مستقیماً با استفاده از یک عبارت منظم‌سازی به مسئله بهینه‌سازی اضافه می‌شود. در نهایت، زیر مسئله درجه دوم (18) را به دست می‌آوریم:

|  |  |
| --- | --- |
| (18) |  |

جایی که و در (11) و (12) تعریف شده اند. حال مجدد در زیر مسئله (18) با در نظر گرفتن پارامتر جریمه قید نابرابری به تابع هدف منتقل می‌گردد و در نتیجه رویکرد جدیدی برای حل زیر مسائل در زمان کمتر براساس تخمین جواب دقیق تر و با سرعت همگرایی بالاتر فراهم می‌آید. بدین منظور زیر مسئله (18) بصورت فرم (19) در نظر گرفته می‌شود:

|  |  |
| --- | --- |
| (19) |  |

**لم1:** مسئله (19) را در نظر بگیرید که در آن پارامتر جریمه و اسکالر مثبت دلخواهی می‌باشد، آنگاه جواب مسئله برابر *است که در آن می‌باشد. علاوه بر آن است و وجود دارد بطوری که جواب بهینه مسئله (18) است اگر*

برقرار باشد و کران بالایی برای می‌باشد.

**اثبات*:*** *ابتدا تابع لاگرانژین مسئله (19) را که بصورت می‌باشد را در نظر می‌گیریم، حال بر اساس شرایط* KKT[[6]](#footnote-6)*داریم:*

|  |  |
| --- | --- |
| *(20)* |  |
| *(21)* |  |

*با استفاده از (20) داریم:*

|  |  |
| --- | --- |
| *(22)* |  |

*حال با استفاده از (21) و (22) داریم:*

|  |  |
| --- | --- |
| *(23)* |  |

*با بکارگیری (23) مقدار محاسبه می‌گردد و در نتیجه:*

|  |  |
| --- | --- |
| *(24)* |  |

*در ادامه نشان می‌دهیم که برای هر کراندار است. در ابتدا از رابطه (24) داریم:*

|  |  |
| --- | --- |
| *(25)* |  |

*با تقسیم رابطه (25) بر داریم:*

|  |  |
| --- | --- |
| *(26)* |  |

*از طرفی می‌دانیم:*

|  |  |
| --- | --- |
| *(27)* |  |

*با حد گرفتن از دو طرف رابطه (26) و بکارگیری (27) داریم:*

|  |  |
| --- | --- |
| *(28)* |  |

*و از (28) نتیجه می‌گیریم:*

|  |  |
| --- | --- |
| *(29)* |  |

*به عنوان یک نتیجه از همگرایی، وجود دارد بطوری که برای هر ،می‌باشد. لذا*

|  |  |
| --- | --- |
|  |  |

*جایی که:*

|  |  |
| --- | --- |
|  |  |

*بوضوح مشخص است اگر*

|  |  |
| --- | --- |
|  |  |

*آنگاه خواهد بود، ازطرف دیگر با فرض ، تابع لاگرانژی (18) به فرم (30) تبدیل می‌شود:*

|  |  |
| --- | --- |
| *(30)* |  |

*و لذا شرایط* KKT *آن بصورت زیر می‌باشند:*

|  |  |
| --- | --- |
|  | *.* |

*بوضوح مشخص است این شرایط برایو برآورده می‌شوند.* ◼

|  |
| --- |
| ***الگوریتم 1:*** *جواب زیر مسئله درجه دوم* |
| ***ورودی:***  ***پارامتر:***  ***مقدار دهی اولیه:***  تکرار مراحل 1تا 4 تا اینکه برقرار شود:    ***4:***  ***خروجی:*** |

*الگوریتم 2 جزئیات روش پیشنهادی را نشان می‌دهد. حلقه بیرونی با هر مقدار اولیه که در آن یک ثابت مثبت دلخواه است، مقدار دهی اولیه می‌شود. سپس، دنباله کاهشی از مقادیر تا زمانی که به آستانه برسند، ایجاد می‌شود ( این پارامتر دقت تقریب های نُرم صفر را کنترل می‌کند). طول گام توسط الگوریتم 1 در طول تکرارها محاسبه می‌شود.*

*مشخص است که انتخاب دقیق پارامتر در (17) می‌تواند ایجاد قید را تضمین کند. از آنجایی که این قید باید فقط در تکرارهای نهایی ایجاد شود، از مقدار کمی شروع و به تدریج در طول تکرارها آن را افزایش می‌دهیم. به این ترتیب، روش را مجبور می‌کنیم در تکرار اولیه الگوریتم تابع هدف را بیشتر از نقض قیود کاهش دهد.*

|  |
| --- |
| ***الگوریتم 2:*** *روش پیشنهادی* (RMSQP) |
| ***ورودی:***  ***پارامترها:***  ***مقدار دهی اولیه:***  *محاسبه مقادیر با کمک روابط (11) و (12) و الگوریتم1*  *خاتمه الگوریتم*  *محاسبه مقادیر با کمک روابط (13) و (14)*    ***خروجی:*** |

*در این بخش تحلیل همگرایی روش پیشنهادی ارائه می‌شود. اثبات همگرایی روش مبتنی بر نتایج بدست آمده توسط فلچر و همکاران* [9] *است و در این خصوص مفروضات استاندارد زیر در نظر گرفته می‌شود:*

1. *تمام نقاطی که توسط الگوریتم تولید می‌شوند در ناحیه بسته و کراندار غیرتهی قرار دارند.*
2. *توابع و در مجموعه باز شامل دوبار مشتق پذیر پیوسته هستند.*
3. *یک وجود دارد به طوری که ماتریس هسین در رابطه صدق می‌کند.*

*یک نتیجه از فرضیات استاندارد این است که ماتریس‌های هسین و روی کراندار هستند و طبق فرض٢، روابط برای هر برقرار است.*

*همگرایی سراسری روش پیشنهادی، با توجه به شرط* KKT*، تحت شرایط قیدی* MF[[7]](#footnote-7)*ثابت می‌شود. نقطه موجه در شرایط قیدی* MF*صدق می‌کند اگر و تنها اگر:*

* *بردارهای برای هر مستقل خطی هستند، که در آن مجموعه به ترتیب شامل مجموعه اندیس برای قیود برابری و نابرابری می‌باشند.*
* *بردار وجود دارد بطوری که*

*که در آن یعنی شامل قیدهای نابرابری که در فعال هستند.*

*شرط لازم برای اینکه نقطه مسئله را حل کند این است که اولا شدنی باشد و ثانیا علاوه بر برقراری شرایط قیدی* MF *مجموعه تهی باشد.*

**قضیه1:** *اگر مفروضات استاندارد برقرار باشد، آنگاه با فرض ثابت بودن یکی از حالات زیر اتفاق می‌افتد:*

*(الف) یک نقطه* KKT *برای مسئله (9) بدست می‌آید. یعنی زیر مسئله را حل می‌کند.*

*(ب) یک نقطه ی تجمعی برای دنباله بدست می‌آید که موجه است. این نقطه* KKT *است و یا در شرایط قیدی* MF *صدق نمی‌کند.*

**اثبات:** *حالت (الف) بوضوح برقرار است. لذا فقط حالت (ب) را اثبات می‌کنیم. ابتدا تقریب سری تیلور تابع را بصورت در نظر می‌گیریم که در آن y نقطه ای روی پاره خط تا می‌باشد. از (13) و (14) نتیجه می‌شود که*

|  |  |
| --- | --- |
| *(31)* |  |

*مقدار را به عبارت (31) اضافه و کم می‌کنیم:*

|  |  |
| --- | --- |
| *(32)* |  |

*براساس داریم:*

|  |  |
| --- | --- |
| *(33)* |  |

*در نتیجه با بکارگیری روابط (32) و (33) داریم:*

|  |  |
| --- | --- |
| *(34)* |  |

*اکنون بر اساس برهان خلف فرض می‌کنیم که یک نقطه تجمع وجود دارد که شرایط قیدی* MF *را برآورده اما در شرایط* KKT *صدق نمی‌کند.*

*بر اساس لم 5 از فلچر و همکاران [9]، همسایگی* *از و ثابت های مثبت وجود دارند، به گونه ای که برای همه و زیرمسئله درجه دوم جواب شدنی دارد که:*

|  |  |
| --- | --- |
| *(35)* |  |

*حال با تقسیم عبارت (34) بر و بکارگیری رابطه (35) داریم:*

|  |  |
| --- | --- |
|  |  |

*و بطور معادل داریم:*

|  |  |
| --- | --- |
|  |  |

*لذا اگر باشد، آنگاه و در نتیجه دنباله مقادیر تابع بطور یکنواخت کاهشی می‌باشد. بنابراین روی مجموعه کراندار و در نتیجه دنباله همگرا خواهد بود و به وضوح با این حقیقت در تناقض می‌باشد، بنابراین یک نقطه* KKT *بوده و لذا حالت (ب) برقرار است.* ◼

1. نتایج عددی

*در این بخش عملکرد عددی الگوریتم‌ پیشنهادی و کارایی آن در بازیابی بردارهای تُنک نشان داده می‌شود. ابتدا تاثیر پارامترهای گوناگون بر روی عملکرد روش بررسی و مقادیر اولیه کارآمدی برای آن‌ها پیشنهاد می‌گردد. در ادامه نتایج مقایسه عددی با برخی از الگوریتم‌های شناخته شده گزارش می‌شود.*

*در آزمایش‌های‌ انجام شده، از بردارهای تُنک و ماتریس حسگر که بطور تصادفی ایجاد شده‌اند استفاده می‌شود و موقعیت مؤلفه‌های غیرصفر از مجموعه { n,...,1} بصورت تصادفی نمونه برداری می‌گردند که در آن است و در نهایت در نظرگرفته می‌شود. از طرفی برابر تقریب جواب بازیابی شده توسط الگوریتم‌ها در نظر گرفته شده است. شاخص عملکرد نرخ سیگنال به نویز*[[8]](#footnote-8)‎ *بصورت زیر تعریف می‌گردد:*

|  |  |
| --- | --- |
|  |  |

***آزمایش اول:*** *در این آزمایش تأثیر پارامترهای , و در کارایی روش پیشنهادی بررسی و برای یافتن مقادیر اولیه کارآمد، شاخص SNR مورد استفاده قرار می گیرد.*

***1) تاثیر پارامتر :*** *اگر مقدار پارامتر کم در نظر گرفته شود، احتمال گیر افتادن در جواب‌های موضعی زیاد می‌شود. از طرف دیگر، انتخاب نسبتاً بزرگ برای باعث هموارسازی بیشتر در تابع SL0 می‌شود و احتمال رسیدن به جواب سراسری را بیشتر می‌کند. شکل 2 میانگین شاخص SNRبعنوان تابعی از c برای ماتریس حسگر در ابعاد 150 ×50 و 200 × 100 را نشان می‌دهد. آزمایش 500 بار تکرار و برای های گوناگون مقدار SNR محاسبه شده است. وقتی کوچک است، همانطور که انتظار می رود شاخص همیشه بزرگ است و از سوی دیگر، برای های به اندازه کافی بزرگ تقریباً بدون تغییر باقی می‌ماند. با بررسی نتایج یک مقدار اولیه کارآمد در نظر گرفته می‌شود.*

|  |  |
| --- | --- |
|  |  |

شکل2. میانگین SNR روش پیشنهادی برای ماتریس در ابعاد (سمت راست) و (سمت چپ).

***2) تاثیر پارامتر :*** *در فرمول (14) که نشان دهنده شرایط کاهش کافی است، پارامتر بین صفر و یک تغییر می‌کند. با در نظر گرفتن مقادیر مختلف به عنوان مقدار اولیه کارآمد پیشنهاد می‌شود. آزمایش 500 بار تکرار و میانگین نمودار شاخص SNR بر حسب در شکل 3 برای مقادیر مختلف نشان داده شده است.*

|  |
| --- |
|  |

شکل3. میانگین شاخص SNR برای مقادیر مختلف در ماتریس حسگر با ابعاد .

***3) تاثیر پارامتر :*** *همانطور که در بخش 2 اشاره شد، پارامتر هموارسازی است. هرچه مقدار کوچکتر باشد دقت تقریب نُرم صفر بهتر خواهد شد. پس از 500 بار تکرار آزمایش، بهترین مقدار برای پارامتر بر اساس نتایج عددی ارائه شده می‌باشد. در شکل 4 میانگین نمودار شاخص* SNR *بر حسب برای مقادیر مختلف نشان داده شده است.*

|  |
| --- |
|  |

شکل4. میانگین شاخص SNR برای مقادیر مختلف  در ماتریس حسگر با ابعاد .

*4) تاثیر پارامتر : برای یافتن مقادیر اولیه کارآمد این پارامتر، آزمایش 500 بار تکرار و در شکل 5 میانگین نمودار شاخص SNR بر حسب گزارش شده است. با بررسی نتایج، مقدار به عنوان مقدار اولیه کارآمد معرفی می‌شود.*

|  |
| --- |
|  |

شکل5. میانگین شاخص SNR برای مقادیر مختلف  در ماتریس حسگر با ابعاد .

***آزمایش دوم:*** *در این آزمایش، ابتدا بصورت مختصر ﻣﺘﺪاول‌ترین الگوریتم‌های بازیابی سیگنال‌های تُنک که با روش پیشنهادی مقایسه شده اند معرفی و در ادامه عملکرد عددی و کارایی آن‌ها با یکدیگر مقایسه می‌شوند.*

* :RMSQP *در روش پیشنهادی با استفاده از الگوریتم 2 و مجموعه پارامترهای زیر، مسئله حل می‌شود.*
* [[9]](#footnote-9)‎ :ILTمسئله بهینه‌سازیرا با آستانه توقف حل می‌کند،که در آن ثابت مثبت کوچک و پارامتر منظم سازی است.
* [[10]](#footnote-10)‎LP: این روش مسئله بهینه‌سازی که در آن می‌باشد را حل می‌کند. این روش بر اساس روش تکراری Re-WL1 توسعه یافته و مقدار به عنوان پارامتر کارآمدی معرفی شده است [10].
* :RSL0[[11]](#footnote-11)‎ این روش که در بخش 2 تشریح شد برای حل مسئله از مجموعه پارامترهای زیر استفاده می‌کند ]11[.
* *[[12]](#footnote-12)* :LASSO این روش مسئله بهینه‌سازی را حل می‌کندکه در آن پارامتر منظم سازی و تابع چگالی تجمعی می‌باشند [12]. پارامترهای اولیه روش بصورت زیر است.
* :SCAD[[13]](#footnote-13)‎ این الگوریتم از یک تابع جریمه غیرمحدب بجای نُرم یک برای تضمین تُنک سازی استفاده می‌کند [13]. پارامترهای اولیه مورد استفاده روش می‌باشد.

پارامتر مهمی که در خصوص ارائه کارایی روش پیشنهادی وجود دارد، بررسی وضعیت ابعاد ماتریس حسگر می‌باشد. در این آزمایش نسبت تعداد ستون ها به تعداد سطرهای ماتریس حسگر در دو حالت مختلف با ابعاد 150×50 () و 200× 100 () در نظر گرفته شد و شاخص عملکرد SNR محاسبه و پس از 500 بار تکرار، میانگین آن گزارش شد. شکل 6 نمودار میانگین شاخص SNR بر حسب در ابعاد 150×50 و 200×100را با مقدار نویز ثابت نشان می‌دهد. مشاهده می‌شود که مقدار شاخص SNR روش پیشنهادی از سایر الگوریتم‌ها بیشتر است و کارایی و عملکرد بهتر روش را نشان می‌دهد.

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

**شکل6. مقایسه روش پیشنهادی RMSQPبا نویز برای (تصویر بالا) و (تصویر پایین).**

**آزمایش سوم:** در این آزمایش، سطح پراکندگی ثابت و به ترتیب برابر با10، 20، 30 و40 در نظر می‌گیریم و مقدار نویز از صفر تا تغییر می‌کند. شکل 7 نمودار SNR بر حسب تغییرات نویز را نشان می‌دهد. همان طور که از شکل مشخص است، برای چهار حالت مختلف از سطح پراکندگی S=10,20,30,40 روش پیشنهادیRMSQP از دقت بهتری نسبت به سایر الگوریتم‌ها برخوردار است.

|  |  |
| --- | --- |
|  |  |
| **s=20** | **s=10** |
|  |  |
| **s=40** | **s=30** |

**شکل7. مقایسه روش پیشنهادی RMSQPبا دامنه تغییرات نویز برای ماتریس حسگر .**

1. نتیجه گیری

در این مقاله، ابتدا الگوریتم‌های موجود برای حل مسائل بهینه‌سازی نمایش تُنُک سیگنال بررسی شد. در ادامه با بهرمندی از برنامه‌ریزی درجه دوم متوالی و نُرم صفر هموارشده روشی جدید و کارا برای بازیابی سیگنال‌های تُنک نویزدار پیشنهاد گردید که همگرایی زبرخطی داشته و نتایج همگرایی سراسری آن در قالب قضیه ١ به خوبی نشان داده شده است که راه حل دقیقی برای حل زیر مسئله با هزینه کم ارائه می‌کند. نتایج خروجی آزمایش‌های انجام شده بر روي داده‌ها جهت بازیابی سیگنال تُنُک نویزدار نشان می‌دهد که روش پیشنهادی نسبت به سایر الگوریتم‌های معرفی شده در این حوزه، نرخ موفقیت بالاتری داشته و از نظر کارایی بهتر است. در این خصوص آزمایش دوم نشان داد که روش پیشنهادی از شاخص SNR بهتری برخوردار است، یعنی بازیابی درخصوص سیگنال تنک انجام شده نسبت به سایر روش‌ها بسیار بهتر می‌باشد و از طرفی مشخص شد روش به ابعاد ماتریس حسگر وابسته نیست یعنی زمانی که تعداد ستون‌ها نسبت به تعداد سطرها سه برابر یا دو برابر باشند، تفاوتی در کارایی روش ایجاد نمی‌کند. آزمایش سوم نیز کارایی بسیار بالای روش را در مقایسه با سایر روش ها در حضور تغییرات نویز از تا نشان داد.

به عنوان پژوهش‌های آینده می‌توان از روش پیشنهادی در مساله بازیابی مسایل تنک غیرمنفی، ماتریس‌های رتبه پایین و بازسازی تصاویر در زمینه‌های پزشکی مانند تصاویر شبکیه چشم یا MRI مغزی بهره منده شد. همچنین می‌توان با در نظرگرفتن مساله ­به‌صورت چندلایه و حل هرکدام از آن‌ها با کمک پردازش موازی، زمان اجرای روش و پیچیدگی مسایل در ابعاد بزرگ را کمتر ساخت. ایده دیگر برای کارهای آتی استفاده از از سایر نرم‌ها بجای نرم صفر هموارشده در تابع هدف می‌باشد، همچنین می‌توان بجای بهره مندی از برنامه‌ریزی درجه دوم متوالی، روش فیلتر را مورد بررسی قرار داد.

1. منابع

[1] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” SIAM Rev., vol. 51, no. 1, pp. 34–81, 2009.

[2] R. Liu, M. Shu, and C. Chen, “ECG Signal Denoising and Reconstruction Based on Basis Pursuit,” Applied Sciences 11, no. 4: 1591, 2021.

[3] W. Jinming and L. Haifeng, “Binary sparse signal recovery with binary matching pursuit,” Inverse Problems., vol. 37, no. 6, pp. 14-65, 2021.

[4] C. Xueping Chen, L. Jianzhong and C. Jiandong, “A new result on recovery sparse signals using orthogonal matching pursuit,” Statistical Theory and Related Fields, pp. 1-7, 2022.

[5] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, “A fast approach for overcomplete sparse decomposition based on smoothed L0 norm,” IEEE Trans. Signal Process., vol. 57, no. 1, pp. 289-301, 2009.

[6] M. S. Alamdari, M. Fatemi, A. Ghaffari, “A modified sequential quadratic programming

method for sparse signal recovery problems,” Signal Processing, 2023.

[7] M. S. Alamdari, M. Fatemi, “Presenting a new method to separate fetal heart signals from the mother by using sequential quadratic programming,” Journal of Advanced Mathematical Modeling, vol. 13, no. 1, pp. 153-167, 2023.

[8] M. S. Alamdari, M. Fatemi, A. Ghaffari, “The Recovery of Sparse Signals by Sequential Quadratic Programming Approach,” Journal of Operational Research and Its Applications, vol. 20, no. 3, pp. 19-32, 2023.

[9] R. Fletcher, S. Leyffer, P.L. Toint, “On the global convergence of a filter–SQP algorithm,” SIAM J. Optim. 13 (1), pp. 44–59, 2002.

[10] S. Foucart and M. Lai, “Sparsest solutions of under-determined linear systems via -minimization for ,” Appl. Comput. Harmon. Anal., vol. 26, no. 3, pp. 395-407, 2009.

[11] A. Eftekhari, M. Babaie-Zadeh, C. Jutten, and H. Abrishami Moghaddam, “Robust-SL0 for stable sparse representation in noisy settings,” in IEEE Int. Conf. Acoust. Speech Signal Process., pp. 3433-3436, 2009.

[12] A. Belloni, V. Chernozhukov, and L. Wang, “Square-root lasso: pivotal recovery of sparse signals via conic programming,” Biometrika, vol. 98, no.4, pp. 791806, 2011.

[13] J. Fan and R. Li, “Variable selection via nonconcave penalized likelihood and its oracle properties,” Journal of the American statistical Association, vol. 96, no. 456, pp. 1348-1360, 2001.

Applying sequential quadratic programming and smoothed L0 norm for recovery of noisy sparse signal

Abstract:

Sparse representation has many applications in signal and image processing, including applications in medical image reconstruction, image enhancement and compression, signal separation, array and radar signal processing. This importance has caused the researchers to benefit from a variety of sparse representation method and to use different norms to solve optimization problems. In this article, firstly, different methods of solving the sparse representation are reviewed, and then a new and efficient method is presented to recover noisy sparse signals with the benefit of sequential quadratic programming and smoothed L0 norm. The results of the experiments show the high success rate of the proposed method compared to other sparse representation optimization methods.

Keywords: Optimization, Sequential quadratic programming, Smoothed L0 norm, Sparse signal representation.

1. Basis Pursuit [↑](#footnote-ref-1)
2. Matching Pursuit [↑](#footnote-ref-2)
3. Orthogonal Matching Pursuit [↑](#footnote-ref-3)
4. Smoothed L0 norm [↑](#footnote-ref-4)
5. Sequential Quadratic Programming [↑](#footnote-ref-5)
6. Karush-Kuhn-Tucker [↑](#footnote-ref-6)
7. Mangasarian-Fromovitz [↑](#footnote-ref-7)
8. Signal-to-Noise Ratio [↑](#footnote-ref-8)
9. Iterative Log Thresholding [↑](#footnote-ref-9)
10. LP quasi-norm [↑](#footnote-ref-10)
11. Robust Smoothed norm [↑](#footnote-ref-11)
12. Least Absolute Shrinkage and Selection Operator [↑](#footnote-ref-12)
13. Smoothly Clipped Absolute Deviation [↑](#footnote-ref-13)