

بازیابی سیگنال‌های تُنک با استفاده از رویکرد برنامه‌ریزی درجه دوم متوالی

محمدسعید علمداری^{۱*}، مسعود فاطمی^۲، ابوذر غفاری^۳

۱- دانشجوی دکتری، دانشکده ریاضی کاربردی، دانشگاه خواجه نصیر الدین طوسی، تهران، ایران

۲- دانشیار، دانشکده ریاضی کاربردی، دانشگاه خواجه نصیر الدین طوسی، تهران، ایران

۳- استادیار، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران

رسید مقاله: ۱۵ بهمن ۱۴۰۱

پذیرش مقاله: ۱۶ خرداد ۱۴۰۲

چکیده

در این مقاله با ایده از روش برنامه‌ریزی درجه دوم متوالی و به کارگیری نرم صفر هموارشده در تابع هدف، روش پیشنهادی برای حل مساله یافتن جواب‌های تُنک دستگاه معادلات خطی فرومعیین ارایه می‌شود. این روش رویکرد جدیدی برای حل زیرمسایل درجه دوم به صورت دقیق فراهم می‌آورد که منجر به پیچیدگی کمتر و سادگی حل زیرمسایل درجه دوم می‌گردد. چارچوب روش پیشنهادی با حدس اولیه برای سیگنال‌های تُنک ناشناخته شروع و در هر تکرار برای محاسبه جهت جستجو یک مساله بهینه‌سازی درجه دوم مشخص حل می‌شود. سپس تحلیل نظری روش ارایه و همگرایی فرو معین آن ثابت می‌شود. نتایج به دست آمده از اجرای روش پیشنهادی بر روی ماتریس‌ها در ابعاد مختلف نشان می‌دهد که کارایی روش به ابعاد ماتریس ورودی وابسته نیست. در نهایت نتایج عددی مبتنی بر مقایسه شاخص SNR روش پیشنهادی با جدیدترین الگوریتم‌های بازیابی سیگنال‌های تُنک گزارش می‌شود که کارایی و عملکرد بالای روش را تاکید می‌کند.

کلمات کلیدی: بازیابی سیگنال‌های تُنک، برنامه‌ریزی غیرخطی، نرم صفر هموارشده، برنامه‌ریزی درجه دوم متوالی.

۱ مقدمه

ایده تجزیه تُنک یا یافتن جواب‌های تُنک دستگاه معادلات خطی فرومعیین یکی از مسایل اساسی در پردازش سیگنال و تصویر بوده و روش‌های متنوعی برای آن معرفی شده است. از جمله روش‌های مورد توجه در این حوزه می‌توان به الگوریتم‌های مبتنی بر نرم صفر هموارشده اشاره کرد که در مقایسه با روش‌های مبتنی بر حداقل کردن مستقیم نرم صفر یا نرم یک، دارای سرعت و صحت بیشتری بوده و ویژگی‌های مطلوب پیوستگی، مشتق‌پذیری، سادگی در پیاده‌سازی و رفتار عددی مناسب را داراست، لذا در سال‌های اخیر مورد توجه محققان قرار گرفته است. مساله یافتن جواب‌های تُنک دستگاه معادلات خطی فرومعیین کاربردهای بسیار زیاد و متنوعی دارد.

* عهده‌دار مکاتبات

آدرس الکترونیکی: m.s.alamdari69@gmail.com

درخصوص برخی کاربردهای آن می توان به بازسازی تصاویر پزشکی، تقویت و فشرده سازی تصاویر، جداسازی سیگنال، سنجش فشرده، تشخیص الگو، تصویربرداری رزونانس مغناطیسی، نمونه برداری از سیگنال های آنالوگ، پردازش سیگنال آرایه و رادار اشاره کرد.

تعداد جواب های دستگاه های معادلات خطی فرومعی، در صورت وجود نامتناهی است ولی شرط تُنک بودن پاسخ، منجر به یافتن پاسخ یکتای مساله می شود. تُنک بودن یک بردار بدین معناست که تعداد عناصر غیر صفر بردار نسبت به تعداد کل درایه ها به اندازه کافی کمتر باشد، یعنی نرم صفر آن کوچک باشد. سیستم معادله خطی (۱) را در نظر بگیرید:

$$y = Ax \quad (1)$$

که در آن $A \in \mathbb{R}^{m \times n}$ ماتریس حسگر با $n > m$ و $y \in \mathbb{R}^m$ بردار سیگنال است. هدف اصلی در مساله بازیابی سیگنال، یافتن تُنک ترین بردار ممکن x است که در (۱) صدق کند. برای حل مساله باید معیاری برای سنجش تُنک بودن بردار در نظر گرفته شود تا با استفاده از آن، مساله بهینه سازی حل و پاسخ یکتای مساله مشخص شود. یک مدل ریاضی رایج، مساله حداقل سازی نرم صفر به فرم (۲) می باشد، یعنی

$$\min_x \|x\| \quad s.t. \quad y = Ax \quad (2)$$

که در آن x معرف تعداد درایه های غیر صفر بردار x است. استفاده از این نرم به دلیل ناپیوستگی، مشتق ناپذیری و حساسیت زیاد به نویز برای حل مسایل بهینه سازی مناسب نیست و به همین دلیل محققان در سال های اخیر به صورت تجربی [۱] و نظری [۵-۲] نشان داده اند که می توان نرم یک را جایگزین نرم صفر کرد و مساله بهینه سازی حداقل سازی نرم یک را حل کرد. معیار دیگری که جایگزین نرم صفر بوده و به عنوان تخمینی از آن معرفی می شود، نرم صفر هموار شده می باشد، یعنی

$$\|x\|_{\sigma} := n - \sum_{i=1}^n \exp\left(-\frac{x_i^2}{\sigma^2}\right) \quad (3)$$

الگوریتم های مبتنی بر نرم صفر هموار شده (۳) در مقایسه با روش های مبتنی بر حداقل کردن نرم صفر و یک، علاوه بر سرعت بیشتر دارای ویژگی های مطلوبی از قبیل پیوستگی، مشتق پذیری، سادگی و رفتار عددی مناسب می باشند که در [۶] گزارش شده است. در این مقاله بر اساس نرم صفر هموار شده روش کارآمدی برای حل (۲) پیشنهاد می شود که اصلاحی از روش برنامه ریزی درجه دوم متوالی [۷] می باشد و ویژگی های آن به شرح زیر است:

(۱) خواص همگرایی سراسری مطلوبی دارد.

(۲) از رویکرد جدیدی برای حل زیرمسایل درجه دوم بهره می برد.

(۳) کاهش پیچیدگی محاسباتی مساله را در پی دارد.

ساختار کلی مقاله بدین شرح است: بخش ۲ به مرور روش نرم صفر هموار شده می پردازد. روش برنامه ریزی درجه دوم متوالی در بخش ۳ یادآوری می گردد. روش پیشنهادی و تحلیل نظری آن در بخش ۴ آورده شده

است. بخش ۵ ارزیابی عملکرد روش پیشنهادی در مقایسه با سایر الگوریتم‌های شناخته شده در زمینه بازیابی سیگنال‌های تُنک را نشان می‌دهد و در نهایت نتیجه گیری و پیشنهادها در بخش ۶ آورده شده است.

۲ مروری بر روش نرم صفر هموار شده

روش‌های متنوعی برای یافتن تُنک‌ترین پاسخ به کار برده می‌شوند و در مورد برخی از آنها می‌توان به روش‌های جستجوی تطابق [۸]، جستجوی پایه [۹]، برنامه‌ریزی خطی [۱۰] و روش فریم‌ها [۱۱] اشاره کرد. روش دیگری که در این زمینه مطرح شده، نرم صفر هموار شده (SL0) نام دارد. این روش در سال ۲۰۰۹ توسط مهمانی و همکارانش [۶] معرفی شد و بر اساس قابلیت روش معرفی شده تابع ناپیوسته نرم صفر توسط تابع پیوسته تخمین زده می‌شود. بدین منظور ابتدا تابع گوسی $f_{\sigma}(x_i)$ با میانگین صفر به فرم (۴) در نظر گرفته می‌شود:

$$f_{\sigma}(x_i) = \exp\left(-\frac{x_i^2}{\sigma^2}\right) \quad (4)$$

به وضوح مشخص است که اگر σ خیلی کوچک باشد بر اساس رابطه (۵) داریم:

$$\lim_{\sigma \rightarrow 0} f_{\sigma}(x_i) = \begin{cases} 1 & x_i = 0 \\ 0 & x_i \neq 0 \end{cases} \quad (5)$$

و در نتیجه نرم صفر هموار شده بردار x به صورت رابطه (۶) تعریف می‌گردد:

$$F_{\sigma}(x) := \sum_{i=1}^n f_{\sigma}(x_i) \quad (6)$$

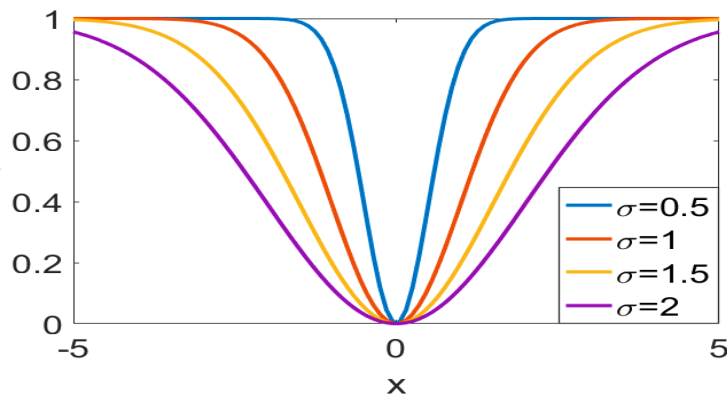
از رابطه (۶) برای مقادیر کوچک σ داریم:

$$\|x\| \approx n - F_{\sigma}(x) \quad (7)$$

حال اگر σ در رابطه (۷) به سمت صفر میل دهیم، تقریب به تساوی تبدیل می‌گردد،

$$\|x\| = n - \lim_{\sigma \rightarrow 0} F_{\sigma}(x) \quad (8)$$

لذا براساس (۸) حداکثر کردن تابع $F_{\sigma}(x)$ برای σ کوچک معادل با حداقل کردن نرم صفر خواهد بود. شکل ۱ نمودار تابع $1 - f_{\sigma}(x)$ را برای مقادیر مختلف x متعلق به $\mathbb{R}$ نشان می‌دهد. می‌توان مشاهده کرد که برای σ کوچک تخمین‌ها به نرم صفر، واقعی‌تر می‌شوند در حالی که هزینه آن ناهموار شدن هر چه بیشتر تابع تقریب بوده که می‌تواند مشکلاتی را در فرایند بهینه‌سازی ایجاد کند از این رو برای جلوگیری از به دام افتادن الگوریتم در بیشینه‌های محلی و کاهش پیچیدگی محاسباتی، مقدار اولیه بزرگی برای این پارامتر در نظر گرفته می‌شود که بتدریج در فرایند تکرارها کاهش می‌یابد.



شکل ۱. نمودار تابع هموار $1 - f_{\sigma}(x)$ برای مقادیر مختلف σ

۳ مروری بر روش برنامه ریزی درجه دوم متوالی

روش برنامه ریزی درجه دوم متوالی^۱ (SQP) یک روش شناخته شده و کارا برای حل مسایل بهینه سازی غیرخطی، غیرمحدب و هموار است. این روش اولین بار توسط ویلسون در سال ۱۹۶۳ 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 مساله بهینه سازی درجه دوم حل می گردد و در آن از تقریب درجه دوم تابع هدف و تقریب خطی درجه اول قیود مساله اصلی استفاده می شود. سرعت همگرایی فرو معین و نتایج همگرایی قوی آن، محققان زیادی را به تحقیقات گسترده در این زمینه سوق داده است [۱۲-۱۴]. مساله بهینه سازی مقید (۹) را در نظر بگیرید:

$$P \begin{cases} \min_x & f_{\sigma}(x) = -\sum_{i=1}^n \exp\left(-\frac{x_i^2}{\sigma^2}\right) \\ \text{s.t.} & c(x) = y - Ax \end{cases} \quad (9)$$

برای محاسبه جهت جدید d زیر مساله ناحیه اعتماد درجه دوم (۱۰) در نظر گرفته می شود:

$$QP(x, \rho) \begin{cases} \min_d & q(d) := g^t d + \frac{1}{2} d^t H d \\ \text{s.t.} & c + \nabla c^t d = 0 \\ & \|d\|_{\rho} \leq \rho \end{cases} \quad (10)$$

که در آن ρ شعاع ناحیه اعتماد، g گرادیان تابع هدف و H هسین تابع می باشد. به صورت دقیق تر داریم:

$$g_i = \frac{2x_i}{\sigma^2} \exp\left(-\frac{x_i^2}{\sigma^2}\right) \quad (11)$$

$$H = \text{Diag}\left(\left(\frac{2}{\sigma^2} - \frac{4x_i^2}{\sigma^4}\right) \exp\left(-\frac{x_i^2}{\sigma^2}\right)\right) \quad (12)$$

که در آن $\text{Diag}(\omega_i)$ ماتریس قطری با عناصر ω_i می باشد. نمادهای Δq به عنوان کاهش پیش بینی شده و Δf به عنوان کاهش واقعی به صورت زیر در نظر گرفته می شوند:

¹ Sequential Quadratic Programming

$$\Delta q := q(\cdot) - q(d) = -g^t d - \frac{1}{2} d^t H d \quad (13)$$

$$\Delta f := f(x) - f(x+d) \quad (14)$$

شرط زیر کاهش کافی در تابع هدف نامیده می شود که در آن $\alpha \in (0, 1)$ است.

$$\Delta f \geq \alpha \Delta q \quad (15)$$

۴ روش پیشنهادی و تحلیل نظری آن

در این مقاله روش پیشنهادی، روش برنامه ریزی درجه دوم متوالی اصلاح شده نامیده می شود. بدین منظور در ابتدا مساله ناهموار (۲) مدنظر می باشد و سپس تقریب همواری از آن به فرم مساله (۹) در نظر گرفته می شود. نقطه شروع الگوریتم نقطه شدنی x صادق در معادله $Ax = y$ است و جستجو در فضای پوچ A صورت می پذیرد و لذا تمامی تکرارها، شدنی باقی می مانند. زیر مساله (۱۰) به فرم (۱۶) تبدیل می شود:

$$QP(x, \rho) \begin{cases} \min_d & g^t d + \frac{1}{2} d^t H d \\ \text{s.t.} & Ad = 0, \\ & \|d\|_2 \leq \rho, \end{cases} \quad (16)$$

که در آن g و H در (۱۱) و (۱۲) تعریف شده اند. لم ۱ درخصوص جواب مساله (۱۶) ارایه شده است که رویکرد جدیدی برای حل زیرمسایل درجه دوم براساس تخمین جواب دقیق در زمان کمتری فراهم می کند. به عبارتی در مساله (۱۶) محدودیت نابرابری $\|d\|_2 \leq \rho$ ، با در نظر گرفتن پارامتر جریمه $\frac{\gamma}{2}$ به تابع هدف منتقل می شود.

لم ۱: مساله زیر را در نظر بگیرید:

$$\min_d \quad g^t d + \frac{1}{2} d^t H d + \frac{\gamma}{2} d^t d \quad \text{s.t.} \quad Ad = 0. \quad (17)$$

که در آن g و H طبق روابط (۱۱) و (۱۲) تعریف شده اند و γ اسکالر مثبت دلخواه می باشد، آنگاه جواب مساله (۱۷) برابر $d^* = -\mathcal{H}^{-1}(A^t \Lambda^* + g)$ است که در آن $\Lambda^* = (-A\mathcal{H}^{-1}A^t)^{-1}(A\mathcal{H}^{-1}g)$ می باشد. علاوه بر آن $\mathcal{H} = H + \gamma I$ است و $0 < \rho_1 \leq \rho$ وجود دارد به طوری که d^* جواب بهینه مساله (۱۶) است اگر رابطه زیر برقرار باشد که در آن M کران بالایی برای Λ^* می باشد.

$$\gamma \geq \frac{\|A\|M + \|g\|}{\rho_1} - \max_{i=1, \dots, n} \left| \left(\frac{2}{\sigma_i^2} - \frac{4x_i^2}{\sigma_i^2} \right) \exp \left(-\frac{x_i^2}{\sigma_i^2} \right) \right|$$

اثبات: ابتدا تابع لاگرانژین مساله (۱۷) را به صورت $\mathcal{L} = \frac{1}{2} d^t \mathcal{H} d + g^t d + \Lambda^t A d$ در نظر می گیریم، حال بر

اساس شرایط KKT داریم:

$$\mathcal{H}d + g + A^t \Lambda = 0. \quad (18)$$

$$Ad = \circ \quad (19)$$

با استفاده از (۱۸) داریم:

$$d^* = -\mathcal{H}^{-1} (A^t \Lambda^* + g) \quad (20)$$

حال با استفاده از (۱۹) و (۲۰) داریم:

$$\circ = Ad = -A\mathcal{H}^{-1} (A^t \Lambda^* + g) \quad (21)$$

با به کارگیری (۲۱) مقدار Λ^* محاسبه می گردد و در نتیجه:

$$\Lambda^* = (-A\mathcal{H}^{-1} A^t)^{-1} A\mathcal{H}^{-1} g \quad (22)$$

در ادامه نشان می دهیم که برای هر $\gamma \geq 0$ ، Λ^* کراندار است. در ابتدا از رابطه (۲۲) داریم:

$$(-A\mathcal{H}^{-1} A^t) \Lambda^* = A\mathcal{H}^{-1} g \quad (23)$$

با تقسیم رابطه (۲۳) بر $\mathcal{H}^{-1}$ داریم:

$$\frac{-A\mathcal{H}^{-1} A^t \Lambda^*}{\|\mathcal{H}\|^{-1}} = A \frac{\mathcal{H}^{-1}}{\|\mathcal{H}\|^{-1}} g \quad (24)$$

از طرفی می دانیم:

$$\lim_{\gamma \rightarrow \infty} \frac{\mathcal{H}^{-1}}{\|\mathcal{H}\|^{-1}} = I \quad (25)$$

با حد گرفتن از دو طرف رابطه (۲۴) و به کارگیری (۲۵) داریم:

$$-AA^t \Lambda^* = Ag \quad (26)$$

و از (۲۶) نتیجه می گیریم:

$$\Lambda^* = (-AA^t)^{-1} Ag \quad (27)$$

به عنوان یک نتیجه از همگرایی، $M > 0$ وجود دارد بطوری که برای هر $\gamma > 0$ ، $\Lambda^* \leq M$ می باشد. لذا

$$\|d\|^* \leq \frac{\|A^t \Lambda^* + g\|}{\|\mathcal{H}\|} \leq \frac{\|A\| M + \|g\|}{\|\mathcal{H}\|}$$

جایی که:

$$\|\mathcal{H}\| = \gamma + \max_{i=1, \dots, n} \left| \left(\frac{2}{\sigma^2} - \frac{4x_i^2}{\sigma^4} \right) \exp \left(-\frac{x_i^2}{\sigma^2} \right) \right|$$

به وضوح مشخص است اگر

$$\frac{\|A\| M + \|g\|}{\rho} - \max_{i=1, \dots, n} \left| \left(\frac{2}{\sigma^2} - \frac{4x_i^2}{\sigma^4} \right) \exp \left(-\frac{x_i^2}{\sigma^2} \right) \right| \leq \gamma$$

آنگاه $\|d\| \leq \rho$ خواهد بود، از طرف دیگر با فرض $\rho_1 = \|d^*\|$ ، تابع لاگرانژی (۱۶) به فرم (۲۸) تبدیل می شود:

$$\mathcal{L} = \frac{1}{\gamma} d^t \mathcal{H} d + g^t d + \Lambda^t A d + \mu (d^t d - \rho_1^2) \quad (28)$$

و لذا شرایط KKT آن به صورت زیر می باشند.

$$\mathcal{H}d + g + A^t\Lambda + \mu d = 0, \quad Ad = 0, \quad d_{\gamma} \leq \rho_{\gamma}, \quad \mu(d^t d - \rho_{\gamma}^2) = 0.$$

به وضوح مشخص است این شرایط برای $\Lambda = \Lambda^*, d = d^*, \mu = 0$ برآورده می شوند. ■

الگوریتم ۱: جواب زیر مساله درجه دوم	
ورودی: A, g, H, γ	
$\mathcal{H} = H + \gamma I$ $\Lambda = (-A\mathcal{H}^{-1}A^t)^{-1}(A\mathcal{H}^{-1}g)$ $d = -\mathcal{H}^{-1}(A^t\Lambda + g)$	خروجی:

الگوریتم ۲ جزئیات روش پیشنهادی را نشان می دهد. حلقه بیرونی با هر مقدار اولیه $\sigma \geq \sigma_{\min}$ که در آن یک ثابت مثبت دلخواه است، مقدار دهی اولیه می شود. سپس، دنباله کاهشی از مقادیر σ تا زمانی که به آستانه $\sigma_{\min}$ برسند، ایجاد می شود که این پارامتر دقت تقریب های نرم صفر را کنترل می کند. شعاع ناحیه اعتماد ρ تا زمانی که (۱۵) برآورده شود، به طور متوالی کاهش می یابد. طول گام d توسط الگوریتم ۱ در طول تکرارها محاسبه می شود. اخیرا در [۱۲] الگوریتم مشابهی برای حل مساله (۲) ارائه شده که در موارد زیر با الگوریتم ۲ تفاوت دارد:

- ۱) کاهش تعداد پارامترهای الگوریتم با تغییر شرط حلقه بیرونی براساس تغییرات σ .
- ۲) کاهش تعداد موارد حل زیر مساله در یک تکرار. در [۱۲] زیر مساله به ازای مقادیر افزایشی پارامتر جریمه تا حصول یک جواب صادق در شرط ناحیه اعتماد حل می شود که این مورد در الگوریتم جدید حذف شده است.
- ۳) بهبود عملکرد روش که نتیجه اقدامات موارد ۱ و ۲ بوده است.

الگوریتم ۲: روش پیشنهادی	
ورودی: $y, A, A^{\dagger} = A^t(AA^t)^{-1}$ $\gamma, c, \alpha, \sigma_{\min}$	
پارامترها: مقدار دهی اولیه:	
$x = A^{\dagger}y, \sigma = 3 \times \max x $ while $\sigma > \sigma_{\min}$ محاسبه مقادیر f, g, H, d if $d = 0$ خاتمه الگوریتم elseif $\Delta f < \alpha \Delta q$ $\gamma \leftarrow \gamma \times 1/2$ end if end while $\sigma \leftarrow c \cdot \sigma$ end while x	خروجی:

در این بخش تحلیل همگرایی روش پیشنهادی ارائه می شود. اثبات همگرایی روش مبتنی بر نتایج به دست آمده

- توسط فلچر و همکاران [۷، ۱۵] است و در این خصوص مفروضات استاندارد زیر در نظر گرفته می شود:
- (۱) تمام نقاطی که توسط الگوریتم تولید می شوند در ناحیه بسته و کراندار غیرتهی X قرار دارند.
 - (۲) توابع $f_\sigma(x)$ و $c(x)$ در مجموعه باز شامل X دوبار مشتق پذیر پیوسته هستند.
 - (۳) یک $M > 0$ وجود دارد به طوری که ماتریس هسین H در رابطه $\|H\|_p \leq M$ صدق می کند.
- یک نتیجه از فرضیات استاندارد این است که ماتریس های هسین $f_\sigma(x)$ و $c(x)$ روی X کراندار هستند و طبق فرض ۲، روابط $\|\nabla^2 c(x)\|_p \leq M$ و $\|\nabla^2 f_\sigma(x)\|_p \leq M$ برای هر $x \in X$ برقرار است.
- همگرایی سراسری روش پیشنهادی، با توجه به شرط KKT ، تحت شرایط قیدی MFCQ ثابت می شود. نقطه موجه x_* در شرایط قیدی MFCQ صدق می کند اگر و تنها اگر:
- بردارهای $\nabla c_i(x_*)$ برای هر $i \in \mathcal{E}$ مستقل خطی است. که مجموعه $\mathcal{E}$ و $\mathcal{I}$ به ترتیب شامل مجموعه اندیس برای قیود تساوی و نامساوی هستند.
 - بردار s وجود دارد به طوری که

$$s^t \nabla c_i(x_*) = 0, i \in \mathcal{E}$$

$$s^t \nabla c_i(x_*) < 0, i \in \mathcal{A}$$

که در آن $\mathcal{A} = \{i \in \mathcal{I} \mid c_i(x_*) = 0\}$ یعنی $\mathcal{A} \subset \mathcal{I}$ شامل قیدهای نامساوی که در x_* فعال هستند.

شرط لازم برای اینکه نقطه x_* مساله را حل کند این است که اولاً شدنی باشد و ثانیاً علاوه بر برقراری قیدی MFCQ مجموعه زیر تهی باشد.

$$\{s \mid s^t g(x_*) < 0\},$$

$$s^t \nabla c_i(x_*) = 0, i \in \mathcal{E},$$

$$s^t \nabla c_i(x_*) < 0, i \in \mathcal{A}\}$$

قضیه ۱: اگر مفروضات استاندارد برقرار باشد، آنگاه با فرض ثابت بودن σ یکی از حالات زیر اتفاق می افتد:

(الف) یک نقطه KKT برای مساله (۹) به دست می آید. یعنی $d = 0$ زیر مساله $QP(x, \rho)$ را حل می کند.

(ب) یک نقطه تجمعی برای دنباله $\{x_k\}$ به دست می آید که موجه است. این نقطه KKT است و یا در شرایط MFCQ صدق نمی کند.

اثبات: حالت (الف) به وضوح برقرار است. لذا فقط حالت (ب) را اثبات می کنیم. ابتدا تقریب سری تیلور تابع

$$f_\sigma(x) \text{ را بصورت } f_\sigma(x_k + d) = f_\sigma(x_k) + g_k^t d + \frac{1}{2} d^t \nabla^2 f(y) d$$

نقطه ای روی پاره خط x_k تا $x_k + d$ می باشد. از (۱۳) و (۱۴) نتیجه می شود که

$$\Delta f_\sigma = -g_k^t d - \frac{1}{2} d^t \nabla^2 f(y) d \quad (29)$$

مقدار $\frac{1}{2} d^t H d$ را به عبارت (۲۹) اضافه و کم می کنیم:

$$\Delta f_{\sigma} = -g_k^t d - \frac{1}{\gamma} d^t \nabla^{\tau} f(y) d + \frac{1}{\gamma} d^t H d - \frac{1}{\gamma} d^t H d \quad (30)$$

$$\Delta f_{\sigma} = \Delta q - \frac{1}{\gamma} d^t (\nabla^{\tau} f(y) - H) d$$

بر اساس $\|d\|_{\gamma} \leq \rho, \|H\|_{\gamma} \leq M, \|\nabla^{\tau} f(y)\|_{\gamma} \leq M$ داریم:

$$\begin{aligned} d^t (\nabla^{\tau} f(y) - H) d &\leq \|d\|_{\gamma} \|\nabla^{\tau} f(y) - H\|_{\gamma} \|d\|_{\gamma} \\ &\leq \|d\|_{\gamma} (\|\nabla^{\tau} f(y)\|_{\gamma} + \|H\|_{\gamma}) \\ &\leq 2M\rho^2 \end{aligned} \quad (31)$$

نتیجه با به کارگیری روابط (۳۰) و (۳۱) داریم:

$$\Delta f_{\sigma} \geq \Delta q - M\rho^2 \quad (32)$$

اکنون بر اساس برهان خلف فرض می کنیم که یک نقطه تجمع x^{∞} وجود دارد که شرایط MFCQ را برآورده اما در شرایط KKT صدق نمی کند.

بر اساس لم ۵ از فلچر و همکاران [۷]، همسایگی $\mathcal{N}^{\infty}$ از x^{∞} و ثابت های مثبت κ, ε وجود دارند، به گونه ای که برای همه $x \in \mathcal{N}^{\infty} \cap X$ و $0 \leq \rho \leq \kappa$ زیرمساله درجه دوم $QP(x, \rho)$ جواب شدنی d دارد که:

$$\Delta q \geq \frac{1}{\gamma} \rho \varepsilon > 0. \quad (32)$$

حال با تقسیم عبارت (۳۲) بر Δq و به کارگیری رابطه (۳۳) داریم:

$$\frac{\Delta f_{\sigma}}{\Delta q} \geq 1 - \frac{\gamma \rho M}{\varepsilon}$$

و به طور معادل داریم:

$$\Delta f_{\sigma} \geq \alpha \Delta q, \quad \alpha = 1 - \frac{\gamma \rho M}{\varepsilon}$$

لذا اگر $\rho \leq (1 - \alpha) \varepsilon / (\gamma M)$ باشد، آنگاه $\Delta f_{\sigma} \geq \alpha \Delta q > 0$ و در نتیجه دنباله مقادیر تابع $f_{\sigma}(x_k)$

به طور یکنواخت کاهشی می باشد. بنابراین $f_{\sigma}(x_k)$ روی مجموعه X کراندار و در نتیجه دنباله $\sum_{k \geq 1} \Delta f_{\sigma}(x_k)$

همگرا خواهد بود و به وضوح با این حقیقت $\Delta f_{\sigma}(x_k) \geq \alpha \Delta q(x_k) \geq \frac{1}{\gamma} \alpha \rho \varepsilon$ در تناقض می باشد، بنابراین x^{∞}

یک نقطه KKT بوده و لذا حالت (ب) برقرار است. ■

۵ ارزیابی عملکرد روش پیشنهادی

در این بخش عملکرد عددی الگوریتم پیشنهادی و کارایی آن در بازیابی بردارهای تُنک نشان داده می شود. ابتدا تاثیر پارامترهای گوناگون بر روی عملکرد روش بررسی و مقادیر اولیه کارآمدی برای آنها پیشنهاد می گردد. در ادامه نتایج مقایسه عددی با جدیدترین الگوریتم های شناخته شده در این زمینه گزارش می شود.

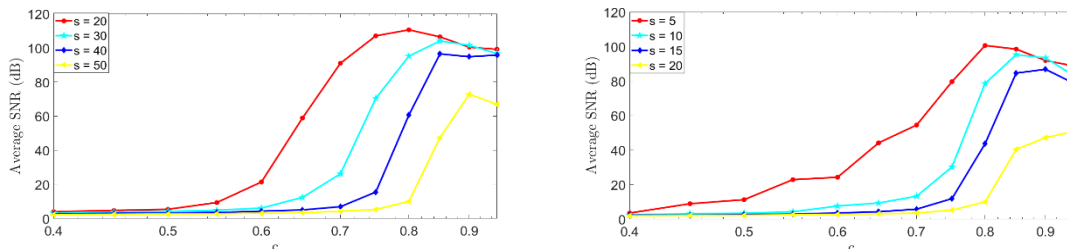
در آزمایش های انجام شده، از بردارهای تُنک و ماتریس حسگر A که به طور تصادفی ایجاد شده اند استفاده

می شود و موقعیت مولفه های غیر صفر از مجموعه $\{1, \dots, n\}$ به صورت تصادفی نمونه برداری می گردند که در آن $\|\bar{x}\| = s$ است و در نهایت $y = A\bar{x}$ در نظر گرفته می شود. از طرفی $\hat{x}$ برابر تقریب جواب بازیابی شده توسط الگوریتم ها در نظر گرفته شده است. شاخص عملکرد نرخ سیگنال به نویز^۱ (SNR) به صورت زیر تعریف می گردد:

$$SNR(\hat{x}, \bar{x}) = 10 \log_{10} \frac{\|\bar{x}\|_2}{\|\bar{x} - \hat{x}\|_2}$$

آزمایش اول: در این آزمایش تاثیر پارامترهای σ_{min}, α, C در کارایی روش پیشنهادی بررسی و برای یافتن مقادیر اولیه کارآمد، شاخص SNR مورد استفاده قرار می گیرد.

(۱) **تاثیر پارامتر C :** اگر مقدار پارامتر C کم در نظر گرفته شود، احتمال گیر افتادن در جواب های موضعی زیاد می شود. از طرف دیگر، انتخاب نسبتاً بزرگ برای C باعث هموارسازی بیشتر در تابع $SL0$ می شود و احتمال رسیدن به جواب سراسری را بیشتر می کند. شکل ۲ میانگین شاخص SNR به عنوان تابعی از C برای ماتریس حسگر در ابعاد 50×150 و 200×100 را نشان می دهد. آزمایش ۵۰۰ بار تکرار و برای S های گوناگون مقدار SNR محاسبه شده است. وقتی S کوچک است، همانطور که انتظار می رود شاخص همیشه بزرگ است و از سوی دیگر، برای S های به اندازه کافی بزرگ تقریباً بدون تغییر باقی می ماند. با بررسی نتایج $C = 0.9$ یک مقدار اولیه کارآمد در نظر گرفته می شود.

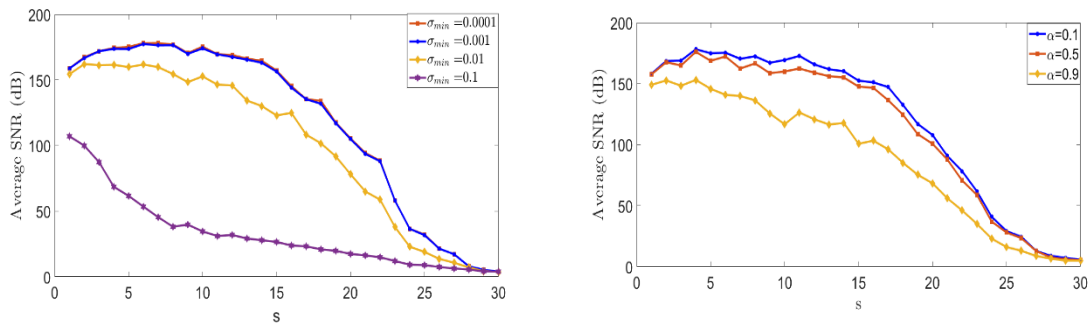


شکل ۲. میانگین شاخص SNR روش پیشنهادی برای ماتریس در ابعاد $A_{50 \times 150}$ (سمت راست) و $A_{100 \times 200}$ (سمت چپ)

(۲) **تاثیر پارامتر α :** در فرمول (۱۵) که نشان دهنده شرایط کاهش کافی است، پارامتر α بین صفر و یک تغییر می کند. با در نظر گرفتن مقادیر مختلف $\alpha = 0.1$ به عنوان مقدار اولیه کارآمد پیشنهاد می شود. آزمایش ۵۰۰ بار تکرار و میانگین نمودار شاخص SNR بر حسب S در شکل ۳ برای مقادیر مختلف α نشان داده شده است.

(۳) **تاثیر پارامتر σ_{min} :** همان طور که در بخش ۲ اشاره شد، σ پارامتر هموارسازی است. هرچه مقدار σ کوچکتر باشد دقت تقریب نرم صفر بهتر خواهد شد. پس از ۵۰۰ بار تکرار آزمایش، بهترین مقدار برای پارامتر σ_{min} بر اساس نتایج عددی ارایه شده $\sigma_{min} = 10^{-3}$ می باشد. در شکل ۳ میانگین نمودار شاخص SNR بر حسب S برای مقادیر مختلف σ_{min} نشان داده شده است.

¹ Signal-to-Noise Ratio



شکل ۳. میانگین شاخص SNR برای مقادیر مختلف α (سمت راست) و مقادیر مختلف σ_{min} (سمت چپ) برای ماتریس با ابعاد $A_{50 \times 150}$

آزمایش دوم:

در این آزمایش، ابتدا به صورت مختصر متداول ترین الگوریتم های بازیابی سیگنال های تُنک که با روش پیشنهادی مقایسه شده اند معرفی و در ادامه عملکرد عددی و کارایی آنها با یکدیگر مقایسه می شود.

- Our Work: در روش پیشنهادی با استفاده از الگوریتم ۲ و پارامترهای زیر، مساله حل می شود.

$$c = 0.9, \alpha = 0.1, \sigma_{min} = 10^{-3}$$

- SL0: این روش که در بخش ۲ تشریح شد برای حل مساله از مجموعه پارامترهای زیر استفاده می کند.

$$c = 0.9, L = 8, \mu = 2, \sigma_{min} = 10^{-4}$$

- Re-WL1: این روش به حل مساله بهینه سازی $\min_x \sum_{i=1}^n \omega_i |x_i|$ s.t. $y = Ax$ می پردازد که در آن $\omega_1, \omega_2, \dots, \omega_n$ وزن های مثبت می باشند [۱۶].

- LP: این روش مساله بهینه سازی $\min_x x_p$ s.t. $y = Ax$ را حل می کند و بر اساس روش تکراری Re-WL1 توسعه یافته است. مقدار $p = 0.5$ به عنوان پارامتر کارآمدی معرفی شده است [۱۷].

- LASSO: این روش مساله بهینه سازی $\min_x \lambda x_1 + Ax - y^2$ را حل می کند که در آن λ پارامتر منظم سازی و Φ تابع چگالی تجمعی می باشند [۱۸].

- SCAD: این الگوریتم از یک تابع جریمه غیر محدب به جای نرم یک برای تضمین تُنک سازی استفاده می کند [۱۹]. پارامترهای اولیه مورد استفاده روش $a = 3/7, \tau = 10^{-3}$ می باشد.

- SEF^۵ و REF^۶: این الگوریتم مساله بهینه سازی $\min_x Ax - y^2 + \lambda g(x)$ را حل می کند [۲۰].

- MMV^۷ و ICDE^۸ و ASDBR^۹ و OMP^{۱۰}: برای مشاهده جزئیات به [۲۱-۲۴] مراجعه نمایید.

¹ Re-Weighted L1

² LP quasi-norm

³ Least Absolute Shrinkage and Selection Operator

⁴ Smoothly Clipped Absolute Deviation

⁵ Shannon Entropy Function

⁶ Renyi Entropy Function

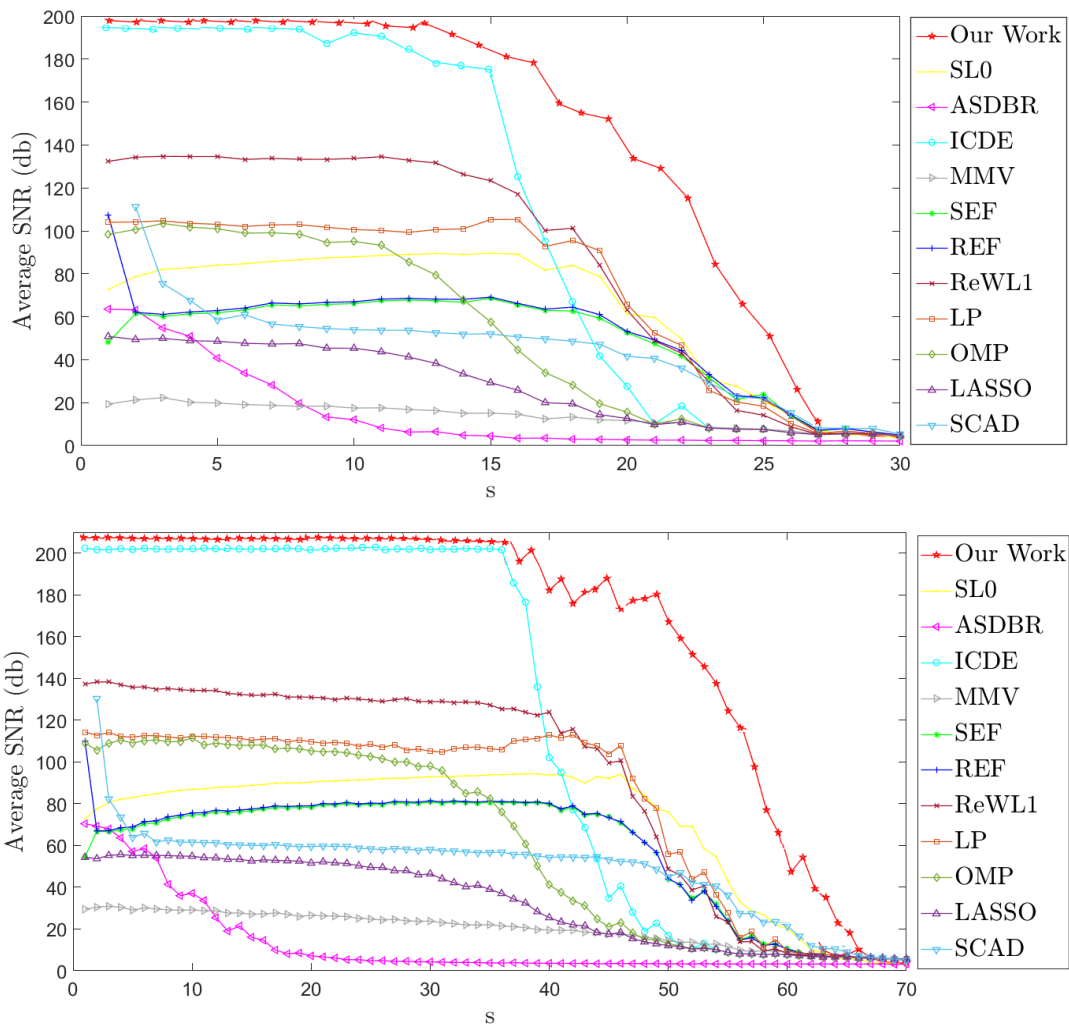
⁷ Multiple Measurement Vector

⁸ Iterative Co-support Detection Estimation

⁹ Adaptive Support Driven Bayesian Reweighted

¹⁰ Orthogonal Matching Pursuit

پارامتر مهمی که درخصوص ارایه کارایی روش پیشنهادی وجود دارد، بررسی وضعیت ابعاد ماتریس حسگر A می باشد. در این آزمایش نسبت تعداد ستون ها به تعداد سطرها ی $(\frac{n}{m})$ ماتریس حسگر $A \in \mathbb{R}^{m \times n}$ در دو حالت مختلف با ابعاد 50×150 ($\frac{n}{m} = 3$) و 100×200 ($\frac{n}{m} = 2$) در نظر گرفته شد و شاخص عملکرد SNR محاسبه و پس از 500 بار تکرار، میانگین آن گزارش شد. شکل 4 نمودار میانگین شاخص SNR بر حسب S برای ماتریس در ابعاد 50×150 و 100×200 را نشان می دهند. در هر دو بخش شکل مشاهده می شود که مقدار شاخص SNR روش پیشنهادی از سایر الگوریتم ها بیشتر است که گویای کارایی و عملکرد بهتر آن می باشد.



شکل 4. مقایسه روش پیشنهادی با سایر روش ها برای ماتریس در ابعاد $A_{50 \times 150}$ (بخش بالا) و $A_{100 \times 200}$ (بخش پایین)

6 نتیجه گیری و پیشنهادها

در این مقاله، روش برنامه ریزی درجه دوم متوالی اصلاح شده برای حل مسایل بهینه سازی بازیابی تُنک پیشنهاد شد که در تابع هدف به جای نرم صفر از نرم هموار شده آن استفاده شد. روش معرفی شده از رویکرد جدیدی برای حل زیرمسایل استفاده می کند و در نتیجه حل کننده کارآمدی معرفی می نماید که منجر به کاهش پیچیدگی محاسبات و ارایه جواب های صریح و دقیق مساله در ابعاد مختلف می گردد.

در مقاله جواب صریح و دقیق زیر مساله در قالب لم ۱ و همگرایی فرو معین آن در قالب قضیه ۱ ارائه می‌شود. برای ارزیابی کار، ابتدا الگوریتم‌های بازیابی سیگنال‌های تُنک معرفی شدند و در ادامه با آزمایش‌های گسترده بر روی مسایل در ابعاد مختلف، نشان داده شد که در همه موارد، روش پیشنهادی از الگوریتم‌های معرفی شده در این زمینه بهتر عمل می‌کند.

در تحقیقات آتی می‌توان موارد زیر را در نظر گرفت:

(۱) در سال‌های اخیر، مساله بازیابی مسایل تُنک غیرمنفی و ماتریس‌های رتبه پایین مورد توجه محققان قرار گرفته است. لذا می‌توان از روش پیشنهادی ارائه شده در این زمینه‌ها بهره‌مند شد.

(۲) بازسازی تصاویر در زمینه‌های پزشکی کاربردهای فراوانی دارد و بسیار مورد توجه است. روش پیشنهادی ارائه شده می‌تواند در بازسازی تصاویر *MRI* مغزی یا شبکه چشم به کارگیری شود.

(۳) در نظر گرفتن مساله به صورت چندلایه و حل هر کدام از آنها با کمک پردازش موازی منجر به کاهش پیچیدگی و زمان اجرای روش در حل مسایل ابعاد بزرگ خواهد شد.

(۴) در این مقاله روش پیشنهادی برای یافتن جواب‌های تُنک، از نرم صفر هموار شده در تابع هدف استفاده می‌نماید که می‌توان به جای آن از سایر نرم‌ها بهره‌مند شد.

(۵) حل مساله ارائه شده در مقاله براساس الگوریتم برنامه‌ریزی درجه دوم متوالی و روش‌های جریمه می‌باشد که می‌توان به جای آن، روش فیلتر را مورد بررسی قرار داد.

منابع

- [1] Malek-Mohammadi, M., Jansson, M., Owrang, A., Koochakzadeh, A., Babaie-Zadeh, M., (2014). DOA estimation in partially correlated noise using low-rank/sparse matrix decomposition. IEEE Sensor Array and Multichannel Signal Processing Workshop, 373-376.
- [2] Donoho, D. L., Elad, M., (2003). Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. Proc of the National Academy of Sciences, 100(5).
- [3] Donoho, D. L., Huo, X., (2001). Uncertainty principles and ideal decomposition. IEEE Trans. Inform. Theory, 47, 2845-2862.
- [4] Elad, M., Bruckstein, A., (2002). A generalized uncertainty principle and sparse representations in pairs or bases. IEEE Trans. Inform. Theory, 48, 2558-2567.
- [5] Donoho, D. L., (2004). For most large underdetermined systems of linear equation the minimal l1 norm solution is also the sparsest solution. Tech. Rep.
- [6] Mohimani, H., Babaie-Zadeh, M., Jutten, C., (2009). A fast approach for overcomplete sparse decomposition based on smoothed l0 norm. IEEE Trans. Signal Process, 57(1), 289-301.
- [7] Fletcher, R., Leyffer, S., Toint, P. L., (2000). On the Global Convergence of a Filter-SQP Algorithm. Technical Report 00/15, Department of Mathematics, University of Namur, Namur, Belgium.
- [8] Jinming, W., Haifeng, L., (2021). Binary sparse signal recovery with binary matching pursuit. Inverse Problems, vol. 37(6), 14-65.
- [9] Liu, R., Shu, M., Chen, C., (2021). ECG Signal Denoising and Reconstruction Based on Basis Pursuit. Applied Sciences, 11(4).
- [10] Wan, A., (2020). Uniform RIP Conditions for Recovery of Sparse Signals by ℓ_p ($0 < p \leq 1$) Minimization. IEEE Transactions on Signal Processing, 68, 5379-5394.
- [11] Chen, S. S., Donoho, D. L., Saunders, M. A., (1999). Atomic decomposition by basis pursuit. SIAM J. Sci. Comput., 20(1), 33-61.
- [12] Alamdari, M. S., Fatemi, M., Ghaffari, A. (2023). A modified sequential quadratic programming method for sparse signal recovery problems. Signal Processing, 207, 108955.

- [13] Bidabadi, N., (2021). Using an Efficient Penalty Method for Solving Linear Least Square Problem with Nonlinear Constraints. *Journal of Operational Research and Its Applications*, 18(4), 21-31. (In Persian)
- [14] Alamdari, M. S., Fatemi, M. (2023). Presenting a new method to separate fetal heart signals from the mother by using sequential quadratic programming. *Journal of Advanced Mathematical Modeling*, 13(1), 153-167. (In Persian)
- [15] Fletcher, R., (2010). The Sequential Quadratic Programming. *Method Nonlinear optimization*, Lecture Note in mathematics, 165-214.
- [16] Candes, E.J., Wakin, M.B., Boyd, S.P., (2008). Enhancing sparsity by reweighted ℓ_1 -minimization. *Journal of Fourier analysis and applications*, 14(5), 877-905.
- [17] Foucart. S., Lai, M., (2009). Sparsest solutions of under-determined linear systems via ℓ_q -minimization for $0 < q \leq 1$. *Appl. Comput. Harmon. Anal.*, 26(3), 395-407.
- [18] Belloni, A., Chernozhukov, V., Wang, L., (2011). Square-root lasso: pivotal recovery of sparse signals via conic programming. *Biometrika*, 98(4), 791806.
- [19] Fan, J., Li, R., (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456), 1348-1360.
- [20] Huang, S., Tran, T.D., (2019). Sparse signal recovery via generalized entropy functions minimization. *IEEE Trans. Signal Process*, 67(5), 1322–1337.
- [21] Song, H., Ren, X. Lai, Y., Meng, H., (2021). Sparse analysis recovery via iterative cosup-port detection estimation, *IEEE Access*, 9, 38386–38395.
- [22] Zonglong, B., Shi, L., Sun. J., Christensen, M., (2020). Complex sparse signal recovery with adaptive laplace priors.
- [23] Li, J., Zhou, W., Cheng, C., (2021). Adaptive support-driven Bayesian reweighted algorithm for sparse signal recovery. *SIViP*, 15, 1295–1302.
- [24] Chen, C.X., Jianzhong, L., Jiandong, C., (2022). A new result on recovery sparse signals using orthogonal matching pursuit. *Stat. Theory Relat. Fields*, 1–7.