

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی اصفهان
دانشکده علوم ریاضی

یک روش شبه نیوتنی گرادیان مزدوج کارا به کمک راه اندازی مجدد و کاربردهای آن

پروژه کارشناسی ریاضی کاربردی

زهرا طهماسبیان مهوار

استاد راهنما

دکتر رضا مختاری

فهرست مطالب

دو	فهرست مطالب
سه	چکیده
۱	۱ مفاهیم مقدماتی
۱	۱.۱ آنالیز تابعی
۲	۲.۱ مفاهیم بهینه‌سازی
۵	۳.۱ همگرایی
۶	۴.۱ روش‌های گرادیان مزدوج
۷	۵.۱ مفاهیم پردازش تصویر
۸	۲ روش شبه‌نیوتنی گرادیان مزدوج
۹	۱.۲ مقدمه
۱۴	۲.۲ الگوریتم طیفی و انگیزه آن
۱۷	۳.۲ تحلیل همگرایی
۲۱	۱.۳.۲ تحلیل همگرایی برای توابع عمومی
۲۵	۴.۲ نتایج عددی
۳۰	۵.۲ کاربردهای GSHS در حل مسائل نوسازی تصویر
۳۴	نتیجه‌گیری
۳۵	کتاب‌نامه
۴۱	واژه‌نامه

چکیده:

برای بهبود ساختار نظری روش گرادیان مزدوج هستینس و استیفل (HS)، اصلاحات مختلفی ارائه شده است که بیشتر آنها بر پایه ویژگی‌های دوگانه برش یافته برای تحلیل همگرایی این روش است. در این گزارش، یک نسخه طیفی از روش HS بررسی شده است که با استفاده از تکنیک راه‌اندازی مجدد پاول، میزان کاهش را به حد مطلوبی می‌رساند و با این اطمینان که روش به صورت کلی همگرا شود. این اصلاحیه امکان کاهش خاصیت کران‌دار بودن در نسخه‌های قبلی روش HS را فراهم می‌کند. علاوه بر این، پارامتر طیفی با توجه به برخی ویژگی‌های نظری مهم در شرط هم‌ارزی تعمیم‌یافته و همچنین خاصیت همگرایی درجه دوم روش نیوتن معرفی شده است. نتایج عددی به دست آمده از مسائل آزمایشی استاندارد نشان می‌دهد که روش پیشنهادی در مقایسه با برخی روش‌های رایج گرادیان مزدوج عملکرد بهتری دارد. همچنین، این روش نتایج قابل اعتمادی را ارائه می‌دهد که در حل مدل‌های بازسازی تصویر کاربرد دارند. کلمات کلیدی: بهینه‌سازی غیرساختاری، روش گرادیان مزدوج طیفی، شرط هم‌ارزی تعمیم‌یافته، جهت نیوتن، همگرایی سراسری، بازسازی تصویر.

فصل ۱

مفاهیم مقدماتی

در این فصل، مفاهیم اولیه و پایه‌ای موردنیاز برای درک روش‌های بهینه‌سازی غیرخطی بررسی می‌شود. این مفاهیم شامل تعاریف اولیه، ویژگی‌های توابع و شرایط لازم برای بهینه‌سازی است. هدف از این بخش، ارائه بستری نظری برای مطالعه روش‌های پیشرفته‌تر در فصل بعدی است.

۱.۱ آنالیز تابعی

قضیه ۱-۱ (نامساوی کوشی-شوارتز). برای هر دو بردار دلخواه $x, y \in X$ در فضای ضرب داخلی X ، نامساوی زیر برقرار است

$$|(x, y)| = |(y, x)| \leq (x, x)^{\frac{1}{2}} (y, y)^{\frac{1}{2}}$$

و برای $y \neq 0$ برابری برقرار می‌شود اگر و تنها اگر به ازای $\alpha \in \mathbb{C}$ داشته باشیم $x = \alpha y$.

تعریف ۱-۲ (نرم-۲ (نرم اقلیدسی)). نرم-۲ یا نرم اقلیدسی یک بردار x در فضای برداری $\mathbb{R}^n$ (یا $\mathbb{C}^n$) به صورت زیر تعریف می‌شود

$$\|x\|_2 = \sqrt{(x, x)}$$

که در آن (x, x) ضرب داخلی استاندارد است یعنی،

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

این نرم، فاصله اقلیدسی یک بردار از مبدأ را نمایش می‌دهد.

قضیه ۱-۳ (تابع به‌طور پیوسته مشتق‌پذیر). فرض کنید $f : X \rightarrow \mathbb{R}$ یک تابع تعریف‌شده روی مجموعه باز $X \subseteq \mathbb{R}^n$ باشد. اگر f در X مشتق‌پذیر باشد و مشتقات جزئی آن در X پیوسته باشند، آنگاه f را به‌طور پیوسته مشتق‌پذیر می‌نامیم و به صورت $f \in C^1(X)$ می‌نویسیم. به‌طور کلی، اگر همه مشتقات جزئی تا مرتبه k در X موجود و پیوسته باشند، آنگاه f را در $C^k(X)$ قرار می‌دهیم.

تعریف ۱-۴ (تابع کراندار). تابع $f : X \rightarrow \mathbb{R}$ را کراندار می‌نامیم اگر عددی مانند $M > 0$ وجود داشته باشد، به‌طوری که برای هر $x \in X$ داشته باشیم $|f(x)| \leq M$.

اگر تابع هم از بالا و هم از پایین کراندار باشد، یعنی دو عدد M_1, M_2 وجود داشته باشند به‌طوری که $M_1 \leq f(x) \leq M_2 \quad \forall x \in X$ ، آنگاه می‌گوییم f دوطرفه کراندار است.

تعریف ۱-۵ (تابع نزولی یکنواخت). تابع $f : X \rightarrow \mathbb{R}$ را نزولی یکنواخت (یا اکیداً نزولی) در مجموعه X می‌نامیم، اگر برای هر دو نقطه $x_1, x_2 \in X$ که $x_1 < x_2$ داشته باشیم $f(x_1) > f(x_2)$.

تعریف ۱-۶ (پیوستگی لیپ‌شیتز). یک تابع $f : \mathbb{R}^n \rightarrow \mathbb{R}$ را لیپ‌شیتز می‌نامند اگر عدد ثابت و مثبت $L > 0$ وجود داشته باشد به‌طوری که برای هر $x, y \in \mathbb{R}^n$ ، رابطه زیر برقرار باشد

$$\|f(x) - f(y)\| \leq L\|x - y\|.$$

اگر این شرط برای گرادیان تابع برقرار باشد، یعنی اگر تابع مشتق‌پذیر باشد و

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y \in \mathbb{R}^n,$$

آنگاه گفته می‌شود گرادیان تابع f لیپ‌شیتز است.

۲۰۱ مفاهیم بهینه‌سازی

تعریف ۱-۲ (محدب شدید). تابع f دوبار مشتق‌پذیر محدب شدید نامیده می‌شود اگر ثابت $m > 0$ و همسایگی حول نقطه مینیم x^* ($\mathcal{N}(x^*)$) چنان وجود داشته باشد که ماتریس هسیان $\nabla^2 f(x)$ در همه نقاط $x \in \mathbb{R}^n$

معین مثبت باشد و شرط زیر برقرار باشد $x \in \mathcal{N}(x^*)$, $\forall v \in \mathbb{R}^n$, $v^T \nabla^2 f(x) v \geq m \|v\|^2$,

تعریف ۸-۱ (معادله سکانتی). در الگوریتم‌های بهینه‌سازی بدون مشتق دوم، برای تقریب هسیان $\nabla^2 f(x)$ ، از معادله سکانتی $B_{k+1} s_k = y_k$ استفاده می‌شود که در آن B_{k+1} تقریب $(k+1)$ -ام ماتریس هسیان، s_k اندازه گام و y_k تغییر در گرادیان است.

تعریف ۹-۱ (جهت جستجوی دقیق). جهت $d_k \in \mathbb{R}^n$ که در تکرار k از الگوریتم بهینه‌سازی مسئله $f(x_k + \alpha d_k)$ را نسبت به $\alpha > 0$ دقیق حل کند را جهت جستجوی دقیق می‌نامند.

تعریف ۱۰-۱ (جهت جستجوی تقریبی). جهت جستجوی تقریبی، جهتی $d_k \in \mathbb{R}^n$ است که به جای حل دقیق مسئله فرعی کمینه‌سازی یا استفاده از ماتریس هسین کامل، از تقریبی از هسین یا اطلاعات گرادیان و جهت‌های قبلی برای تولید جهت کاهش استفاده می‌کند.

تعریف ۱۱-۱ (جهت جستجوی بهینه). جهت جستجوی بهینه، جهتی $d_k \in \mathbb{R}^n$ است که در یک معیار خاص، حداکثر کاهش ممکن را در تابع هدف $f(x)$ در یک گام فراهم کند.

تعریف ۱۲-۱ (جهت نیوتن). نقطه فعلی x_k از دنباله تولیدشده توسط الگوریتم تکراری را در نظر بگیرید. جهت نیوتن d_k به صورت زیر تعریف می‌شود

$$d_k = -\nabla f(x_k) [\nabla^2 f(x_k)]^{-1}.$$

تعریف ۱۳-۱ (شرط محدودشده دوطرفه). تابع f ، جهت جستجو d_k و پارامتر طیفی ξ_k را در نظر بگیرید. گوئیم پارامتر طیفی ξ_k در شرط محدودشده دوطرفه زیر صدق می‌کند

$$\eta_1 \leq \xi_k \leq \eta_2,$$

که در آن $\eta_1 > 0$ و $\eta_2 > 0$ ثابت‌هایی از پیش تعیین شده هستند. این شرط تضمین می‌کند که مقدار ξ_k نه آن قدر کوچک باشد که باعث کندی همگرایی الگوریتم شود، و نه آن قدر بزرگ که موجب واگرایی یا ناپایداری روش گردد. به عبارت دیگر، کران‌دار بودن پارامتر طیفی نقش مهمی در پایداری عددی و عملکرد مؤثر الگوریتم‌های بهینه‌سازی ایفا می‌کند. انتخاب مناسب بازه $[\eta_1, \eta_2]$ معمولاً با آزمون تجربی یا تحلیل نظری مسئله تعیین می‌شود.

تعریف ۱۴-۱ (شرط هم‌ارزی تعمیم‌یافته). در یک الگوریتم بهینه‌سازی که دنباله‌ای از نقاط $\{x_k\}$ را به منظور حل مسئله

$$\min_{x \in \mathbb{R}^n} f(x)$$

تولید می‌کند، شرط هم‌ارزی تعمیم‌یافته معمولاً به صورت یک ویژگی یا معیار تعریف می‌شود که نشان می‌دهد جهت جستجو d_k یا اندازه گام α_k الگوریتم، به‌طور تقریبی با جهت یا گام متناظر در یک روش مرجع (مانند گرادیان کاهشی یا نیوتن) منطبق است. این شرط به‌منظور تضمین رفتار مشابه با الگوریتم‌های کلاسیک و حفظ ویژگی‌های همگرایی آن‌ها استفاده می‌شود.

تعریف ۱-۱۵ (شرط ولف قوی). در بهینه‌سازی، شرط ولف قوی مجموعه‌ای از معیارها برای انتخاب طول گام در روش‌های جستجوی خطی است که تضمین می‌کند الگوریتم به سمت کاهش تابع هدف حرکت کند و در عین حال از گام‌های بیش از حد کوچک یا بزرگ جلوگیری شود. این شرط شامل دو بخش است

(۱) کاهش کافی، برای یک تابع هدف $f(x)$ ، طول گام α_k باید به گونه‌ای انتخاب شود که مقدار تابع در نقطه جدید به اندازه کافی کاهش یابد

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k \nabla f(x_k)^T d_k,$$

که در آن $0 < c_1 < 1$ یک پارامتر ثابت است (معمولاً مقداری کوچک مثل 10^{-4})، $\nabla f(x_k)$ گرادیان تابع در نقطه x_k و d_k جهت جستجو است.

(۲) شرط انحنای قوی، گرادیان تابع در نقطه جدید باید به اندازه کافی در جهت مخالف جهت جستجو باشد تا از گام‌های بیش از حد کوچک جلوگیری شود

$$|\nabla f(x_k + \alpha_k d_k)^T d_k| \leq c_2 |\nabla f(x_k)^T d_k|,$$

که در آن $0 < c_1 < c_2 < 1$ (معمولاً c_2 مقداری نزدیک به ۱، مثل ۰/۹) است. این شرط تضمین می‌کند که گرادیان در جهت d_k به اندازه کافی کوچک شود یا تغییر علامت دهد.

تعریف ۱-۱۶ (پدیده جمینگ (گیرکردگی)). در بهینه‌سازی، پدیده جمینگ (گیرکردگی) به حالتی اشاره دارد که یک الگوریتم تکراری به دلیل ویژگی‌های تابع هدف $f(x)$ یا محدودیت‌های الگوریتم، در یک نقطه x_k متوقف می‌شود یا پیشرفت آن بسیار کند می‌شود. این حالات عبارتند از کمینه محلی، الگوریتم در یک کمینه محلی گیر می‌کند، جایی که $\nabla f(x_k) = 0$ است، اما x_k کمینه سراسری نیست. نقاط زینی، الگوریتم در نقاطی که گرادیان صفر است اما هسین نه مثبت معین و نه منفی معین است، متوقف می‌شود. گرادیان‌های کوچک، در مناطق مسطح تابع (جایی که $\|\nabla f(x_k)\|$ بسیار کوچک است)، الگوریتم به کندی حرکت می‌کند.

۳.۱ همگرایی

برای یک الگوریتم بهینه‌سازی که دنباله‌ای از نقاط را برای کمینه‌سازی تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ تولید می‌کند مفاهیم زیر را تعریف می‌کنیم

تعریف ۱-۱۷ (همگرایی سراسری). روش را همگرای سراسری می‌نامیم، اگر برای هر نقطه شروع اولیه $x_0 \in \mathbb{R}^n$ در مجموعه سطح $S_0 = \{x \in \mathbb{R}^n : f(x) < f(x_0)\}$ الگوریتم به یک نقطه ثابت یا به یک مقدار بهینه همگرا شود به عبارت دیگر، اگر $\{x_k\}$ دنباله تولیدشده توسط روش عددی و x^* جواب دقیق آن باشد، آنگاه

$$\lim_{k \rightarrow \infty} x_k = x^*.$$

تعریف ۱-۱۸ (ویژگی قوی همگرایی سراسری). یک روش عددی دارای ویژگی قوی همگرایی سراسری است، اگر برای هر مقدار اولیه دلخواه x_0 ، دنباله تولیدشده توسط روش، با نرخ همگرایی مشخص و یکنواخت به جواب دقیق همگرا شود. به عبارت دیگر، اگر x_k دنباله تولیدشده توسط روش عددی و x^* جواب دقیق باشد، آنگاه وجود یک نرخ همگرایی یکنواخت C و ρ با $0 < \rho < 1$ داریم، به طوری که

$$\|x_{k+1} - x^*\| \leq C \|x_k - x^*\|^\rho, \quad \forall k \geq 0.$$

تعریف ۱-۱۹ (ویژگی همگرایی درجه دوم). فرض کنید e_k خطا در مرحله k -ام و $C > 0$ یک مقدار ثابت باشد آنگاه

$$\lim_{k \rightarrow \infty} \frac{|e_{k+1}|}{|e_k|^2} = C.$$

این رابطه نشان می‌دهد که اگر مقدار خطا در یک گام کوچک باشد، در گام بعدی خطا با نرخ درجه دوم کاهش می‌یابد. به عبارت دیگر، هر بار که به جواب دقیق نزدیک‌تر می‌شویم کاهش خطا به صورت شتاب‌دار اتفاق می‌افتد.

تعریف ۱-۲۰ (خاصیت نزول کافی). جهت جستجو d_k که الگوریتم در هر مرحله تولید می‌کند اگر در شرط زیر صدق کند گوئیم خاصیت نزول کافی است، که ρ یک عدد ثابت مثبت است

$$g_k^T d_k \leq -\rho \|g_k\|^2.$$

تعریف ۱-۲۱ (ویژگی همگرایی پایدار). گوئیم الگوریتم تکراری برای یافتن مینیمم تابع f دارای همگرایی پایدار است هرگاه دو خاصیت زیر را برآورده کند.

۱ دارای همگرایی سراسری

۲ خاصیت نزول کافی

تعریف ۱-۲۲ (شرط همگرایی مطلوب روش نیوتن). فرض کنید تابع f دوبار مشتق‌پذیر باشد. گوییم روش نیوتن با تکرار $x_{k+1} = x_k - \alpha_k [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$ تحت شرط همگرایی مطلوب قرار دارد اگر

۱ تابع f در همسایگی نقطه بهینه x^* ($\nabla f(x^*) = 0$) به شدت محدب باشد.

۲ ماتریس هسیان $\nabla^2 f(x)$ در همسایگی x^* پیوستگی لپ‌شیتز دارد.

۳ نقطه اولیه x_0 به اندازه کافی به نقطه بهینه x^* نزدیک باشد.

۴ اندازه گام $\alpha_k > 0$ شرایط خط جستجو مناسب، مانند شرایط ولف قوی، را برآورده کند.

تحت این شرایط روش نیوتن دارای همگرایی درجه دوم است.

۴.۱ روش‌های گرادیان مزدوج

در این بخش قصد داریم یک نمای کلی از روش گرادیان مزدوج و مشتقات آن داشته باشیم تا برای خواننده ابهامی در طول متن وجود نداشته باشد ابتدا مروری بر تعریف روش گرادیان مزدوج خواهیم داشت و سپس به حالات مختلف آن اشاره می‌کنیم.

تعریف ۱-۲۳ (روش گرادیان مزدوج). برای مسئله بهینه‌سازی غیرساختاری $\min_{x \in \mathbb{R}^n} f(x)$ ، که f حداقل یک‌بار مشتق‌پذیر است، روش گرادیان مزدوج به صورت تکراری نقاط x_k را با استفاده از رابطه زیر به‌روزرسانی می‌کند که در آن جهت جستجو است که به صورت $d_k = -g_k$ ، $d_{k+1} = -g_{k+1} + \beta_k d_k$ ، تعریف می‌شود، با $g_k = \nabla f(x_k)$ (گرادیان در x_k)، β_k پارامتری که جهت‌های مزدوج را تضمین می‌کند و $\alpha_k > 0$ اندازه گام است

$$x_{k+1} = x_k + \alpha_k d_k.$$

در فصل ۲ روش گرادیان مزدوج طیفی، روش گرادیان مزدوج غیرخطی، روش گرادیان مزدوج ترکیبی، روش HS طیفی سه جمله‌ای توضیح داده می‌شود.

۵.۱ مفاهیم پردازش تصویر

تعریف ۱-۲۴ (مجموعه نویزهای کاندید). در مسائل بازسازی تصویر، داده مشاهده شده (y) معمولاً به صورت زیر مدل می شود

$$y = Ax + \epsilon$$

که در آن x تصویر یا سیگنال اصلی (ناشناخته)، A عملگر مشاهده (مثلاً ماتریس تاری یا نمونه برداری) و ϵ نویز (اغتشاش تصادفی یا سیستماتیک) است. مجموعه نویزهای کاندید (N) مجموعه ای از مدل های نویز یا نمونه های نویزی (ϵ_i) است که برای شبیه سازی یا تخمین نویز در داده های مشاهده شده استفاده می شود.

تعریف ۱-۲۵ (فیلتر میانگین تطبیقی). فرض کنید $I(x, y)$ مقدار پیکسل در موقعیت (x, y) یک تصویر باشد و هدف، محاسبه مقدار فیلترشده $\hat{I}(x, y)$ باشد. فیلتر میانگین تطبیقی معمولاً به صورت زیر عمل می کند

۱ انتخاب پنجره محلی، یک پنجره W با اندازه $m \times m$ حول پیکسل (x, y) تعریف می شود.

۲ محاسبه ویژگی های محلی

میانگین محلی، $\mu_W = \frac{1}{|W|} \sum_{(i,j) \in W} I(i, j)$ که $|W|$ تعداد پیکسل های پنجره است.

واریانس محلی، $\sigma_W^2 = \frac{1}{|W|} \sum_{(i,j) \in W} (I(i, j) - \mu_W)^2$

واریانس نویز تخمینی (σ_n^2) ، معمولاً از کل تصویر یا به صورت پیش فرض تعیین می شود.

۳ محاسبه مقدار فیلترشده، $\hat{I}(x, y) = \mu_W + \frac{\max(0, \sigma_W^2 - \sigma_n^2)}{\sigma_W^2} (I(x, y) - \mu_W)$

اگر $\sigma_W^2 \approx \sigma_n^2$ (منطقه هموار با نویز بالا)، مقدار فیلترشده به μ_W نزدیک می شود، یعنی فیلتر مانند میانگین استاندارد عمل می کند. اگر $\sigma_W^2 \gg \sigma_n^2$ (منطقه با جزئیات یا لبه ها)، مقدار فیلترشده به $I(x, y)$ نزدیک می شود، یعنی تصویر اصلی حفظ می شود. $\max(0, \sigma_W^2 - \sigma_n^2)$ از منفی شدن وزن جلوگیری می کند.

تعریف ۱-۲۶ (تکنیک پروفایل عملکرد). فرض کنید مجموعه ای از مسائل آزمایشی $P = \{1, 2, \dots, n_p\}$ و مجموعه ای از الگوریتم ها (حل کننده ها) $S = \{1, 2, \dots, n_s\}$ داریم. برای هر مسئله $p \in P$ و هر الگوریتم $s \in S$ ، معیاری از عملکرد (مانند زمان محاسباتی یا تعداد ارزیابی های تابع) به صورت $t_{p,s}$ تعریف می شود. پروفایل عملکرد برای الگوریتم s به صورت تابع تجمعی $\rho_s(\tau) = \frac{1}{n_p} |\{p \in P : t_{p,s} \leq \tau\}|$ تعریف می شود.

فصل ۲

روش شبه‌نیوتنی گرادیان مزدوج

در این فصل، یکی از روش‌های مهم و کارآمد در بهینه‌سازی غیرخطی، یعنی روش مزدوج گرادیان شبه‌نیوتنی مورد بررسی قرار می‌گیرد. این روش با ترکیب ویژگی‌های روش‌های گرادیان و نیوتنی، توانسته است کارایی بالایی را در حل مسائل ابعاد بزرگ ارائه دهد. از جمله ویژگی‌های بارز این روش می‌توان به کاهش نیاز به محاسبه مستقیم ماتریس‌های هسیان و مصرف کمتر حافظه اشاره کرد که آن را به گزینه‌ای مناسب برای کاربردهای عملی تبدیل می‌کند.

در ادامه، فرمول‌بندی دقیق این روش ارائه شده است و ویژگی‌های همگرایی آن بررسی می‌شود. همچنین، برخی از اصلاحیه‌های پیشنهادی بررسی خواهند شد که بهبود عملکرد این روش را در شرایط مختلف تضمین می‌کنند. این تحلیل به خوانندگان کمک می‌کند تا درک بهتری از نحوه عملکرد روش گرادیان مزدوج شبه‌نیوتنی داشته باشند و بتوانند آن را در مسائل مختلف بهینه‌سازی به کار گیرند.

۱.۲ مقدمه

این مطالعه بر روی فرمول‌های تکراری گرادیان مزدوج غیرخطی برای حل کردن مسئله بهینه‌سازی غیرساختاری تمرکز دارد که در آن $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ تابعی از پایین کراندار و گرادیان $g(x) = \nabla f(x)$ موجود است،

$$\min f(x), x \in \mathbb{R}. \quad (۱.۲)$$

فرمول گرادیان مزدوج غیرخطی به علت کاربردهای بی‌شمار آن در دنیای واقعی از جمله، انتخاب پرتفوی (انتخاب سبد سرمایه‌گذاری) [۳۱]، حرکت رباتیک [۴۲، ۴۹]، بازیابی سیگنال [۲۱] و مسائل بازسازی تصویر [۴۳] برای اصلاحیه در نظر گرفته شده است. این روش به دلیل ویژگی‌های قوی همگرایی سراسری، نیاز کم به حافظه و کارایی عددی بالا در حل مسائل بهینه‌سازی در مقیاس بزرگ، به عنوان یک الگوریتم تکراری به طور گسترده شناخته شده و معروف برای حل معادله (۱.۲) محسوب می‌شود. با یک حدس اولیه $x^0 \in \mathbb{R}^n$ شروع می‌شود، الگوریتم گرادیان مزدوج دنباله‌ای از تکرارها را به فرم $\{x_k\}$ به وسیله فرمول زیر تولید می‌کند

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \dots, \quad (۲.۲)$$

و α_k اندازه گام است که برای یافتن طول گام مناسب به کار برده می‌شود و باید شرایط استاندارد ولف زیر را برآورده کند

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (۳.۲)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k. \quad (۴.۲)$$

شرط انحنای تعریف شده در نامساوی (۴.۲) افزایش کافی گرادیان را تضمین می‌کند. با این حال، ایراد اصلی روش‌های ولف (استاندارد) این است که مقدار اندازه گام α_k ، ممکن است مقدار نزدیکی به مقدار کمینه‌کننده تابع هدف نداشته باشد که این موضوع اغلب بر نتایج همگرایی تأثیر می‌گذارد. این شرط (۴.۲) به صورت

$$|g(x_k + \alpha_k d_k)^T d_k| \leq -\sigma g_k^T d_k, \quad (۵.۲)$$

تعدیل شده است با ترکیب رابطه (۳.۲) با شرط انحنای تعدیل شده (۵.۲)، جست‌وجوی خطی ولف قوی برقرار می‌شود، با شرط $0 < \delta < \frac{1}{4}, \delta < \sigma < 1$. یکی دیگر از اجزای مهم روش گرادیان مزدوج غیرخطی، جهت

جستجوی d_k است که با روابط

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad d_1 = -g_1 \quad (۶.۲)$$

محاسبه می‌شود. در اینجا پارامتر اسکالر β_k ، به عنوان ضریب گرادیان مزدوج عمل می‌کند که فرمول‌های مختلف گرادیان مزدوج را متمایز می‌سازد. یک بررسی جامع از این موضوع توسط هیگر و جانگ در بخش [۲۰] ارائه شده‌است. فرمول‌های کلاسیک و معروف روش گرادیان مزدوج به دو گروه تقسیم می‌شوند. گروه اول شامل فرمول‌های هسیتنس و استیفل (HS) [۲۲]، پولاک، ربیرو و پالیاک (PRP) [۳۵، ۳۶] و لو و استوری (LS) [۳۰] است که در آن‌ها انتخاب‌های مختلف پارامتر β_k به شکل‌های مختلف مشخص شده‌است،

$$\beta_k^{HS} = \frac{g_{k+1} y_k}{d_k^T y_k}, \quad \beta_k^{PRP} = \frac{g_{k+1} y_k}{\|g_k\|^2}, \quad \beta_k^{LS} = -\frac{g_{k+1} y_k}{d_k^T g_k} \quad (۷.۲)$$

که $\| \cdot \|$ بیانگر نرم l_2 (نرم-۲) و $y_k = g_{k+1} - g_k$ است. واضح است که عبارت موجود در صورت کسر در تمامی سه فرمول یکسان است در حالی که عبارت موجود در مخرج کسر، این روش‌ها را از یکدیگر متمایز می‌سازد [۳]. در محاسبات عملی، فرمول‌های گرادیان مزدوج غیرخطی HS، PRP و LS به‌طور کلی به عنوان روش‌هایی با عملکرد عددی کارآمد شناخته می‌شوند. با این حال، همگرایی روش PRP همیشه برای توابع عمومی تضمین نشده‌است و روش HS ویژگی نزولی را ندارد [۲۰].

گروه دوم که شامل روش‌هایی مانند فلتچر و رویز (FR) [۱۶]، دایه و یوان (DY) [۱۱] و کاهشی مزدوج (CD) [۱۵] می‌شوند با فرمول‌های

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}, \quad \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}, \quad \beta_k^{CD} = -\frac{\|g_{k+1}\|^2}{d_k^T g_k} \quad (۸.۲)$$

مشخص می‌شوند. این روش‌ها دارای ویژگی‌های همگرایی پایدار هستند. با این حال، عملکرد عددی آنها ممکن است تحت تاثیر پدیده‌های جمینگ (گیرکردگی) قرار گیرد [۲۹]. این معایب منجر به اصلاحات متعددی در فرمول‌های گرادیان مزدوج شده‌است به [۲، ۲۳، ۲۸، ۳۹، ۴۰، ۵۱] مراجعه کنید.

تلاش‌های زیادی صرف کشف اصلاحات بهبودیافته فرمول‌های گرادیان مزدوج شده‌است که نه تنها جهتی کافی و عملکرد محاسباتی کارآمد را فراهم می‌کنند، بلکه برخی شرایط هم‌ارزی را نیز برآورده می‌سازند. با گنجاندن پارامترهای مختلف روش گرادیان مزدوج، چندین محقق توانسته‌اند ویژگی‌های همگرایی و کارایی محاسباتی

این الگوریتم‌های جدید را به‌طور مؤثری بهبود بخشند. به عنوان مثال، با در نظر گرفتن شرط هم‌ارزی

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k, \quad (9.2)$$

دایه و لیاو [۱۰] فرمول جدیدی برای روش گرادیان مزدوج بررسی کردند

$$\beta_k^{DL} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - t \frac{g_{k+1}^T s_k}{d_k^T y_k}, \quad (10.2)$$

که $s_k = x_{k+1} - x_k$. برای تضمین همگرایی سراسری برای توابع عمومی، فرمول (۱۰.۲) را به شکل زیر برای بعضی پارامتر t تنظیم کرده‌اند

$$\beta_k^{DL*} = \max\left\{\frac{g_{k+1}^T y_k}{d_k^T y_k}, 0\right\} - t \frac{g_{k+1}^T s_k}{d_k^T y_k}. \quad (11.2)$$

واضح است که وقتی $t = 0$ ، معادله (۱۰.۲) به روش کلاسیک گرادیان مزدوج HS تبدیل می‌شود. با گسترش این مفهوم به روش BFGS بدون حافظه و خودمقیاس‌دهی، هیگر و جانگ [۱۸] فرمول دیگری به فرم

$$\beta_k^N = \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\|y_k\|^2}{(d_k^T y_k)^2} g_{k+1}^T d_k \quad (12.2)$$

بررسی کرده‌اند و نشان دادند که فرمول (۱۲.۲) شرط نزولی $\frac{\gamma}{\lambda} \|g_{k+1}\|^2 \leq -g_{k+1}^T d_{k+1}$ را برآورده می‌کند. برای تضمین همگرایی سراسری برای توابع عمومی، آنها نسخه محدود شده‌ای از فرمول (۱۲.۲) را به صورت $\beta_k^{N+} = \max\left\{\beta_k^N, \frac{-1}{\|d_k\| \min\{\eta, \|g_k\|\}}\right\}$ ، $\eta > 0$ ارائه داده‌اند. نتایج حاصل از محاسبات عددی نشان داده‌اند که این روش کارآمد و امیدوارکننده است. برای یک بررسی جامع از پیشرفت‌های اخیر روش‌های دایه و لیاو به مقاله بابایی و کافاکی [۴] مراجعه کنید.

اخیراً، محققان به روش گرادیان مزدوج ترکیبی پرداخته‌اند، که یک تعمیم اساسی و موثر از روش گرادیان مزدوج غیرخطی است و تحت شرایط مناسب شرط نزول کافی را برآورده می‌کند. این فرمول معمولاً با ترکیب فرمول‌های سایر روش‌های گرادیان مزدوج موجود و کلاسیک به دست می‌آید. یکی از جنبه‌های اساسی و چالش برانگیز الگوریتم گرادیان مزدوج ترکیبی شامل انتخاب پارامتر گرادیان مزدوج β_k و جهت جستجو d_k است. این دو فرمول نقش مهمی در تحلیل همگرایی و عملکرد محاسباتی هر الگوریتم ترکیبی ایفا می‌کنند. لو و لی [۲۹] یک پارامتر جدید روش گرادیان مزدوج را بررسی می‌کنند که می‌توان آن را به عنوان ترکیب محدب فرمول‌های DY و LS در نظر گرفت $\beta_k = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{DY}$ ، که در آن β_k^{LS} از رابطه (۷.۲) و β_k^{DY} از رابطه (۸.۲)

به دست می‌آیند و θ_k به صورت $\theta_k = \frac{g_{k+1}^T y_k d_k^T g_{k+1} - s_k^T g_{k+1} d_k^T y_k}{\|g_{k+1}\|^2 d_k^T g_k + g_{k+1}^T y_k d_k^T y_k}$ تعریف می‌شود. نویسندگان نشان داده‌اند که جهت جست‌وجوی d_k که توسط فرمول آنها تولید می‌شود شرایط هم‌ارزی معروف D-L ارائه‌شده توسط لو و لی [۱۰] را برآورده می‌کند و سپس ویژگی نزول کافی و تحلیل همگرایی را تحت شرایط ولف قوی مورد بحث قرار دادند. نتایج حاصل از آزمایش‌های محاسباتی نشان‌دهنده مزیت این روش بود.

روش‌های گرادیان مزدوج طیفی تغییرات اضافی در جهت جست‌وجوی گرادیان مزدوج ارائه می‌دهند تا ساختار نظری روش‌های گرادیان مزدوج را بهبود بخشند. این روش‌ها بر اساس کارهای ریدن [۳۸]، بارزلای و بورواین [۶] و همچنین برگن و مارتینز [۷] توسعه یافته‌اند که در ابتدا این روش‌ها را پیشنهاد دادند. برای نتایج جدیدتر درباره روش‌های گرادیان مزدوج طیفی می‌توان به منابع [۸، ۳۳، ۴۱، ۴۱، ۴۲، ۴۴، ۴۸] مراجعه کرد. به‌طور مشابه، شرط هم‌ارزی استاندارد $y_k^T d_{k+1} = 0$ نقش مهمی در تحلیل همگرایی روش‌های گرادیان مزدوج خاص ایفا می‌کند [۴۵]. این روش‌ها اغلب به معیارهای جست‌وجوی خطی دقیق برای انتخاب اندازه گام α_k وابسته هستند. به هر حال، این نیاز برای مسائل با ابعاد بزرگ از نظر محاسباتی پرهزینه می‌شود [۲۹]. بنابراین، فرمول‌بندی روش‌های گرادیان مزدوج با جست‌وجوی خطی تقریبی معمولاً برای طراحی الگوریتم‌های موثر ترجیح داده می‌شوند. [۴۵] در صورتی که $y_k^T d_{k+1} \neq 0$ دایه و لیاو [۱۰] پیشنهاد کردند از شرط هم‌ارزی تعمیم‌یافته‌ای که در رابطه (۹.۲) آمده است، استفاده شود.

به عنوان نمونه، وانگ [۴۸] به تازگی یک روش HS طیفی سه‌جمله‌ای را با استفاده از این مفهوم بررسی کرد که به‌طور ذاتی از طریق طراحی خود ویژگی نزول کافی را برآورده می‌کند. با این حال، برخی از اصلاحات در روش‌های HS طیفی به شرط محدودشده دوطرفه ارائه‌شده توسط فرامرزی و امینی در بخش [۱۴] پایبند هستند

$$\theta_k = \begin{cases} \theta_k^M, & \theta_k^M \in [\eta_1, \eta_2], \\ 1, & \text{در غیر این صورت,} \end{cases} \quad (13.2)$$

در این جا η_1 و η_2 کران‌های بالا و پایین مثبت و θ_k پارامتر طیفی به صورت

$$\theta_k^M = \frac{1}{z_{k-1}^T g_k} (s_{k-1}^T g_k + g_k^T z_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}} \|z_{k-1}\|^2). \quad (14.2)$$

است. این اصلاح بر اساس ویژگی‌های نظری طرح شبه‌نیوتنی است که پیش‌تر توسط جیان و همکاران در بخش [۲۴] پیشنهاد شده بود و از پارامترهای طیفی به صورت

$$\theta_k = \begin{cases} \theta_k^{JC\pm}, & \theta_k^{JC\pm} \in [\frac{1}{\tau} + \eta, \tau], \\ 1, & \text{در غیر این صورت,} \end{cases} \quad (15.2)$$

استفاده می‌شود. در این جا η و τ پارامترهای مثبت و θ_k^{J+} به صورت

$$\theta_k^{J+} = 1 - \frac{\|y_k\|^2 d_k^T g_{k+1}}{(d_k^T y_k)(g_{k+1}^T y_k)} + \frac{s_k^T g_{k+1}}{g_{k+1}^T y_k} \quad (۱۶.۲)$$

تعریف می‌شود. برای اطمینان از این که جهت جستجو شرط نزول کافی را برآورده می‌کند، یعنی

$$d_{k+1}^T g_{k+1} \leq -\rho \|g_{k+1}\|^2, \quad \forall k \geq 0 \quad (۱۷.۲)$$

آنها برای دستیابی به فرمول نهایی جمله آخر را حذف کردند و به رابطه زیر رسیدند

$$\theta_k^{J-} = 1 - \frac{\|y_k\|^2 d_k^T g - k + 1}{(d_k^T y_k)(g_{k+1}^T y_k)}, \quad (۱۸.۲)$$

به عبارتی پارامترهای موجود در روابط (۱۴.۲) و (۱۶.۲) و (۱۸.۲) شرایط رابطه (۱۷.۲) را برآورده کرده‌اند و با استفاده از کران‌های بالا و پایین در روابط (۱۵.۲) و (۱۳.۲) به صورت سراسری همگرا می‌شوند. نتایج مشابهی درباره همگرایی اصلاحات روش HS نیز توسط چن و همکاران [۸] و همچنین خوشسیمای بارگرد و اشرفی [۲۶] ارائه شده‌است.

سوال: آیا امکان دارد این اصلاحات روش HS بدون اعمال ویژگی کران‌دار بودن همچنان رابطه (۱۷.۲) را برآورده کنند و به‌طور سراسری همگرا شوند؟

با الهام از ساختار مناسب روش گرادیان مزدوج ترکیبی در بخش [۲۹]، استفاده از ایده جیان و همکاران در بخش [۲۴]، در نظر گرفتن شرط هم‌ارزی تعمیم‌یافته و همچنین همگرایی مطلوب روش نیوتن یک روش اصلاح‌شده گرادیان مزدوج HS برای مسائل با ابعاد بزرگ بهینه‌سازی غیرساختاری در نظر گرفته شده‌است. این روش دارای ویژگی نزول کافی است و با استفاده از تکنیک راه‌اندازی مجدد پاول به صورت سراسری همگرا می‌شود. این اصلاح باعث کاهش شرط محدودشده دوطرفه اعمال‌شده توسط جیان و همکاران [۲۴]، فرامرزی و امینی [۱۴] و اخیراً امینی و فرامرزی [۱] در نسخه‌های قبلی اصلاحات روش HS می‌شود. بر اساس بعضی مسائل آزمایشی استاندارد، نتایج عددی نشان می‌دهد که این روش در مقایسه با برخی روش‌های گرادیان مزدوج دیگر عملکرد امیدوارکننده‌ای دارد. علاوه بر این، روش پیشنهادی هنگام حل مدل‌های بازسازی تصویر نیز قابلیت اطمینان بالایی نشان داده‌است. بخش بعدی به فرآیند استخراج روش طیفی ساختاریافته اختصاص دارد و در ادامه در بخش ۳ تحلیل همگرایی ارائه می‌شود. در بخش ۴، عملکرد پارامتر پیشنهادی را بر روی برخی مسائل معیار بررسی می‌شود و با سایر روش‌های مشابه مقایسه شده‌است. در نهایت، در بخش پایانی برخی نتیجه‌گیری‌ها ارائه شده‌است.

۲.۲ الگوریتم طیفی و انگیزه آن

برای کاهش ویژگی کران دار بودن در فرمول بندی روش های HS قبلی، می توان جهت جستجو d_k در رابطه (۶.۲) را با یک پارامتر طیفی ξ_{k+1} مقیاس بندی کرد

$$d_{k+1} = -\xi_{k+1}g_{k+1} + \beta_k^H S d_k, \quad \forall k \geq 0. \quad (۱۹.۲)$$

این مقیاس بندی تضمین می کند که اگر نقطه فعلی x_{k+1} به اندازه کافی به یک کمینه کننده محلی نزدیک باشد و تابع هدف در بعضی همسایگی ها مانند یک تابع درجه دوم عمل کند، آنگاه جهت جستجوی بهینه همان جهت نیوتن خواهد بود

$$d_{k+1} = -\nabla^2 f(x_{k+1})^{-1} g_{k+1}. \quad (۲۰.۲)$$

که در آن $\nabla^2 f(x_{k+1}) \in \mathbb{R}^{n \times n}$ ماتریس هسیان f در نقطه x_{k+1} است. روش نیوتن برای بروزرسانی رابطه (۲۰.۲) نیاز به ماتریس هسیان، یعنی اطلاعات مشتق دوم، دارد که نرخ همگرایی مطلوبی را فراهم می کند. با الهام از ویژگی همگرایی درجه دوم روش نیوتن، فرض شود که $\nabla^2 f(x_{k+1})$ در هر تکرار وجود دارد و یک معادله سکانتی مناسب را برآورده می کند. برای مثال

$$\nabla^2 f(x_{k+1}) s_k = y_K \quad (۲۱.۲)$$

با برابر قرار دادن رابطه (۱۹.۲) با رابطه (۲۰.۲) داریم

$$-\nabla^2 f(x_{k+1})^{-1} g_{k+1} = -\xi_{k+1}g_{k+1} + \beta_k^{HS} d_k. \quad (۲۲.۲)$$

این رابطه همراه با رابطه (۲۱.۲) نشان می دهد که روش نیوتن طیفی (SNM) دارای ساختار زیر است

$$\begin{aligned} -s_k^T g_{k+1} &= -\xi_{k+1}^{SNM} s_k^T \nabla^2 f(x_{k+1}) g_{k+1} + \beta_k^{HS} s_k^T \nabla^2 f(x_{k+1}) d_k. \\ -s_k^T g_{k+1} &= -\xi_{k+1}^{SNM} y_k^T g_{k+1} + \beta_k^{HS} y_k^T d_k. \end{aligned}$$

بعد از ساده سازی واضح می شود که $\xi_{k+1}^{SNM} = \frac{s_k^T g_{k+1}}{y_k^T g_{k+1}} + \beta_k^{HS} \frac{y_k^T d_k}{y_k^T g_{k+1}}$ با اعمال β_k^{HS} از رابطه (۷.۲) نتیجه به صورت $\xi_{k+1}^{SNM} = \frac{s_k^T g_{k+1}}{y_k^T g_{k+1}} + 1$ خواهد بود. با توجه به اینکه هر الگوریتم گرادیان مزدوج با ساختار نیوتنی

دارای یک خاصیت محلی است، برای فرمول‌بندی یک الگوریتم با خاصیت سراسری، به جای رابطه (۲۰.۲) از شرط هم‌ارزی معروف D-L استفاده می‌شود. با ضرب اولیه رابطه (۱۹.۲) در y_k^T و ترکیب آن با رابطه (۹.۲)، ساختار طیفی دایه و لیاو (SDL) به صورت $-ts_k^T g_{k+1} = -\xi_{k+1}^{SDL} y_k^T g_{k+1} + \beta_k^{HS} y_k^T d_k$ ، به دست می‌آید. یعنی

$$\xi_{k+1}^{SDL} = \frac{ts_k^T g_{k+1}}{y_k^T g_{k+1}} + \beta_k^{HS} \frac{y_k^T d_k}{y_k^T g_{k+1}}. \quad (23.2)$$

پس از انجام برخی ساده‌سازی‌ها، فرمول (۲۳.۲) به شکل زیر خواهد بود

$$\xi_{k+1}^{SDL} = \frac{ts_k^T g_{k+1}}{y_k^T g_{k+1}} + 1. \quad (24.2)$$

تذکر ۱-۲. از رابطه (۲۴.۲) مشاهده می‌شود که وقتی $t = 1$ باشد پارامتر $\xi_{k+1}^{SDL} = \xi_{k+1}^{SNM}$. این یعنی جهت جست‌وجوی طیفی d_{k+1} نه تنها شرط هم‌ارزی تعمیم‌یافته D-L را برآورده می‌کند بلکه از ویژگی همگرایی مطلوب روش نیوتن نیز بهره‌برداری می‌کند. با این حال با انتخاب $t = 1$ تنها نتایج بهتری در همسایگی کمینه‌کننده تضمین می‌شود. در نتیجه، پارامتر باید به صورت خودتنظیمی باشد.

انتخاب پارامتر t را می‌توان به صورت زیر استنتاج کرد

$$t = \frac{y_k^T s_k}{\|s_k\|^2} + \frac{\|y_k\|}{\|s_k\|}. \quad (25.2)$$

این پارامتر t در موقعیت‌های مختلف قابل استفاده است و این انتخاب از [۵، ۹] اتخاذ می‌شود تا همگرایی سراسری روش پیشنهادشده تضمین شود. پس، برای اطمینان از ویژگی نزول کافی، مقدار بهینه HS طیفی (SHS) به صورت زیر محاسبه می‌شود

$$\xi_{k+1}^{SHS} = \max\{1, 1 + t \frac{s_k^T g_{k+1}}{y_k^T g_{k+1}}\}. \quad (26.2)$$

تذکر ۲-۲. باید توجه داشت که این فرمول‌بندی امکان کاهش شرط محدودشده دوطرفه اعمال‌شده در نسخه‌های قبلی اصلاحات HS [۱۴، ۲۴] را فراهم می‌کند همان طور که در روابط (۱۳.۲) و (۱۵.۲) ذکر شده‌است. اکنون روش HS رابطه (۱۷.۲) را برآورده کرده است و بدون حذف جمله حاوی $s_k^T g_{k+1}$ همگرایی سراسری را تضمین می‌کند. ساختار ارائه‌شده در رابطه (۲۶.۲) نشان می‌دهد که $\xi_{k+1} \geq 1$ برقرار است.

حال، الگوریتم جدید SHS به صورت زیر توصیف می‌شود،

الگوریتم ۱ SHS

گام ۱: $x_0 \in \mathbb{R}^n$ و پارامتر $0 < \delta < \sigma < 1$ را انتخاب کن. مقداردهی اولیه $d_0 = -g_0$ و $\alpha_0 = 1$.
 گام ۲: الگوریتم را زمانی که $\|g_k\| \leq \epsilon$ متوقف کن.
 گام ۳: اندازه گام α_k را به گونه‌ای تعیین کن که شرط ولف قوی را ارضا کند.
 گام ۴: پارامتر طیفی ξ_{k+1}^{SHS} با استفاده از رابطه (۲۶.۲) و مقدار t از (۲۵.۲) و β_k^{HS} را از (۷.۲) محاسبه کن.
 گام ۵: جهت جستجو را با رابطه $d = -\xi_{k+1}^{SHS} g_{k+1} + \beta_k^{HS} d_k$ محاسبه کن.
 سپس شرط زیر را بررسی کن

$$|g_{k+1}^T g_k| > \sigma/2 \|g_{k+1}\|^2. \quad (27.2)$$

اگر این شرط برقرار بود $d_{k+1} = -g_{k+1}$ قرار بده در غیر این صورت، $d = d_{k+1}$.
 گام ۶: تکرار جدید را با $x_{k+1} = x_k + \alpha_k d_k$ تولید کن.
 گام ۷: $k = k + 1$ قرار بده و به گام ۲ برو.

باید توجه داشت از آنجا که در الگوریتم SHS $q_k \neq 0$ و با توجه به تعریف ξ_{k+1} در رابطه (۲۶.۲)، مقدار $\xi_{k+1} \neq 0$ خواهد بود. این امر تضمین می‌کند که ξ_{k+1} به خوبی تعریف شده‌است. توجه داشته باشید که با استفاده از روابط (۵.۲) و (۴.۲) داریم

$$\begin{aligned} d_k^T y_k &= d_k^T g_{k+1} - d_k^T g_k \\ &\geq -(1 - \sigma) d_k^T g_k \\ &\geq \frac{(1 - \sigma)}{\sigma} |g_{k+1}^T d_k|. \end{aligned}$$

یعنی نامساوی زیر دراثبات ویژگی نزول کافی روش طیفی مفید خواهد بود

$$\frac{|g_{k+1}^T d_k|}{|d_k^T y_k|} \leq \frac{\sigma}{1 - \sigma}. \quad (28.2)$$

قضیه ۲-۳. فرض کنید ξ_{k+1} طبق رابطه (۲۶.۲) داده شده باشد و الگوریتم SHS دنباله‌های $\{d_k\}$ و $\{g_k\}$ را تولید کند. آنگاه، یک ثابت $\rho > 0$ وجود دارد به‌طوری که شرط زیر برقرار است

$$d_{k+1}^T g_{k+1} \leq -\rho \|g_{k+1}\|^2, \quad \forall k \geq 0. \quad (29.2)$$

معیار (۲۹.۲) تضمین می‌کند که جهت (۱۹.۲) به‌طور مؤثری برای دستیابی به کمینه کافی است.

اثبات. فرض کنید که معیار راه‌اندازی مجدد پاول [۳۱] در شرط (۲۷.۲) برقرار نباشد. پس، نامساوی زیر برقرار خواهد بود

$$|g_{k+1}^T g_k| \leq \frac{1}{2} \|g_{k+1}\|^2, \quad (30.2)$$

چون $y_k = g_{k+1} - g_k$ با استفاده از رابطه (۳۰.۲) به دست می‌آید،

$$|g_{k+1}^T y_k| \leq \|g_{k+1}\|^2 + |g_{k+1}^T g_k| \leq \|g_{k+1}\|^2 + \frac{1}{2} \|g_{k+1}\|^2 = \frac{3}{2} \|g_{k+1}\|^2. \quad (31.2)$$

برای به دست آوردن یک روش گرادیان مزدوج که شرط (۲۹.۲) را برآورده کند، از رابطه (۱۹.۲) و g_{k+1}^T استفاده می‌شود تا به نتیجه

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\xi_{k+1} \|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{d_k^T y_k} g_{k+1}^T d_k \\ &\leq -\xi_{k+1} \|g_{k+1}\|^2 + \frac{|g_{k+1}^T d_k|}{|d_k^T y_k|} |g_{k+1}^T y_k|, \end{aligned}$$

دست یابد. با اعمال نامساوی (۲۸.۲) و (۳۱.۲) در نامساوی اخیر،

$$g_{k+1}^T d_{k+1} \leq -\xi_{k+1} \|g_{k+1}\|^2 + \frac{1/2\sigma}{1-\sigma} \|g_{k+1}\|^2 = -(\xi_{k+1} - \frac{1/2\sigma}{1-\sigma}) \|g_{k+1}\|^2. \quad (32.2)$$

برای اثبات ویژگی نزول کافی، عبارت $\xi_{k+1} - \frac{1/2\sigma}{1-\sigma}$ باید مثبت باشد. با قراردادن $\rho = \xi_{k+1} - \frac{1/2\sigma}{1-\sigma}$ و $\sigma \in (0, 0.4]$ این شرط با پارامتر اسکالر $\rho \geq 1$ برآورده می‌شود که باعث تعریف روش SHS در رابطه (۲۶.۲) می‌گردد. $\square$

۳.۲ تحلیل همگرایی

تحلیل همگرایی نیازمند فرضیات زیر است

فرض ۴-۲. با داشتن نقطه اولیه x_0 ، تابع $f(x)$ روی مجموعه سطح $S_0 = \{s \in \mathbb{R}^n : f(x_0) \geq f(s)\}$ از پایین کراندار است، همچنین ثابت مثبت B چنان وجود دارد که $\|x\| \leq B, \quad \forall x \in S_0$.

فرض ۵-۲. در یک همسایگی T از S_0 ، تابع f به‌طور پیوسته مشتق‌پذیر است و گرادیان آن پیوسته لیپ‌شیتس است یعنی، یک ثابت مثبت B چنان وجود دارد که $\|g(x) - g(y)\| \leq B\|x - y\|, \quad \forall x, y \in T$ بر

اساس فرض ۲-۴، ثابت مثبت γ وجود دارد به طوری که $\|g(x)\| \leq \gamma, \quad \forall x \in S$ ، با استفاده از تکنیک جست و جوی خطی (۳.۲) و معیار نزولی کافی (۲۹.۲) واضح است که دنباله $\{f(x_k)\}$ نزولی یکنواخت است. چون f از پایین کراندار است، ثابت f^* چنان وجود دارد که $\lim_{k \rightarrow \infty} f(x_k) = f^*$. بنابراین، لم زیر گرفته شده از [۵۳] برای تحلیل همگرایی ضروری است.

لم ۲-۶. فرض کنید که $\{x_k\}$ و $\{d_k\}$ دنباله های تولید شده توسط الگوریتم گرادیان مزدوج باشند به طوری که جهت جست و جوی d_k به صورت نزولی انتخاب شده باشد و α_k شرط ولف قوی را برآورده کند. در این صورت داریم $\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^4} < \infty$.

توجه داشته باشید α_k که شرط (۵.۲) را برآورده می کند، باید شرط (۴.۲) را نیز برآورده کند. اما عکس این رابطه لزوماً درست نیست، همان طور که در [۲۹] مشاهده شده است. بنابراین، لم ۲-۶ با جست و جوی خطی ولف قوی نیز برقرار است.

اثبات. براساس رابطه (۴.۲) و فرض ۲-۴ داریم

$$(\sigma - 1)d_k^T g_k \leq (g_{k+1} - g_k)^T d_k \leq \|g_{k+1} - g_k\| \|d_k\| \leq L\alpha_k \|d_k\|^2.$$

سپس $\alpha_k \geq \frac{(\sigma-1)}{L} \frac{(d_k^T g_k)}{\|d_k\|^2}$ نتیجه گرفته می شود. با استفاده از رابطه (۳.۲) عبارت زیر را خواهیم داشت $f(x_k) - f(x_k + \alpha_k d_k) \geq \delta \frac{(1-\sigma)}{L} \frac{(d_k^T g_k)^2}{\|d_k\|^4}$ با فرض ۲-۴ و ترکیب نامساوی های اخیر به دست خواهد آمد $\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^4} < \infty$. $\square$

لم ۲-۷. فرض کنید فرض ۲-۴ برقرار باشد و دنباله های $\{x_k\}, \{d_k\}$ توسط الگوریتم SHS با $\xi_{k+1} > \frac{1-\sigma}{1-\sigma}$ تولید شوند و جهت جست و جوی طیفی d_k نزولی باشد و α_k انتخاب شده شرط ولف قوی را برآورده کند، آنگاه فقط یکی از دو عبارت زیر برقرار است

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0 \quad (33.2)$$

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (34.2)$$

اثبات. از آنجایی که $\sigma \in (0, 0.4]$ و $\xi_{k+1} > \frac{1-\sigma}{1-\sigma}$ جهت جست و جوی d_k تولید شده توسط الگوریتم SHS به

ازای هر k نزولی است. یعنی $g_{k+1}^T d_{k+1} < 0$. با اعمال رابطه (۱۹.۲) داریم

$$\begin{aligned} (\beta_k^{HS})^2 \|d_k\|^2 &= \|d_{k+1} + \xi_{k+1} g_{k+1}\|^2 \\ &= \|d_{k+1}\|^2 + 2\xi_{k+1} g_{k+1}^T d_{k+1} + \xi_{k+1}^2 \|g_{k+1}\|^2 \\ &\leq \|d_{k+1}\|^2 + \xi_{k+1}^2 \|g_{k+1}\|^2. \end{aligned}$$

پس نتیجه گرفته می شود

$$\|d_{k+1}\|^2 \geq (\beta_k^{HS})^2 \|d_k\|^2 - \xi_{k+1}^2 \|g_{k+1}\|^2. \quad (۳۵.۲)$$

از طرفی دیگر، چون d_{k+1} یک جهت نزولی است و مطابق با رابطه (۱۹.۲) همچنین خواهیم داشت

$$g_{k+1}^T d_{k+1} - \beta_k^{HS} g_{k+1}^T d_k = -\xi_{k+1} \|g_{k+1}\|^2 \quad \text{یعنی}$$

$$|\beta_k^{HS}| |g_{k+1}^T d_k| + |g_{k+1}^T d_{k+1}| \geq \xi_{k+1} \|g_{k+1}\|^2. \quad (۳۶.۲)$$

علاوه بر این، با در نظر گرفتن حداکثر مقداری که σ می تواند در رابطه (۵.۲) بگیرد زمانی که رابطه (۳۲.۲) با $\zeta_{k+1} = 1$ برقرار است نتیجه می شود که $\sigma_{max} = 0.4$. بنابراین، رابطه (۵.۲) را می توان به صورت

$$|g_{k+1}^T d_k| \leq -0.4 g_k^T d_k = 0.4 |g_k^T d_k| \quad \text{نوشت. ترکیب با رابطه (۳۶.۲) نتیجه می دهد}$$

$$0.4 |\beta_k^{HS}| |g_k^T d_k| + |g_{k+1}^T d_{k+1}| \geq \xi_{k+1} \|g_{k+1}\|^2. \quad (۳۷.۲)$$

باید توجه داشت که، برای هر $a, b \in \mathbb{R}$ نامساوی $(a + 0.4b)^2 \leq (1 + 0.4^2)(a^2 + b^2)$ همیشه برقرار است. با به کارگیری نامساوی بالا با $a = |g_{k+1}^T d_{k+1}|$ و $b = |\beta_k^{HS} g_k^T d_k|$ روی رابطه (۳۷.۲) نتیجه می دهد که

$$c = \xi_{k+1}^2 / (1 + 0.4^2), \quad \text{با قرار دادن } (g_{k+1}^T d_{k+1})^2 + (\beta_k^{HS})^2 (g_k^T d_k)^2 \geq \frac{\xi_{k+1}^2}{1 + 0.4^2} \|g_{k+1}\|^4$$

$$\begin{aligned} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(g_k^T d_k)^2}{\|d_k\|^2} &= \frac{1}{\|d_{k+1}\|^2} \left[(g_{k+1}^T d_{k+1})^2 + \frac{\|d_{k+1}\|^2}{\|d_k\|^2} (g_k^T d_k)^2 \right] \\ &\geq \frac{1}{\|d_{k+1}\|^2} \left[c \|g_{k+1}\|^4 + (g_k^T d_k)^2 \left(\frac{\|d_{k+1}\|^2}{\|d_k\|^2} - (\beta_k^{HS})^2 \right) \right]. \quad (۳۸.۲) \end{aligned}$$

به طور مشابه از رابطه (۳۵.۲) به دست می آید $\frac{\|d_{k+1}\|^2}{\|d_k\|^2} \geq (\beta_k^{HS})^2 - \xi_{k+1}^2 \frac{\|g_{k+1}\|^2}{\|d_k\|^2}$. یعنی

$$\frac{\|d_{k+1}\|^2}{\|d_k\|^2} - (\beta_k^{HS})^2 \geq -\xi_{k+1}^2 \frac{\|g_{k+1}\|^2}{\|d_k\|^2}. \quad (۳۹.۲)$$

رابطه (۳۹.۲) و رابطه (۳۸.۲) عبارت زیر را که $\xi_{k+1} > \frac{1/2\sigma}{1-\sigma}$ نتیجه می دهند

$$\begin{aligned} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(g_k^T d_k)^2}{\|d_k\|^2} &\geq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^4} \left[c - \xi_{k+1}^2 \frac{(g_k^T d_k)^2}{\|d_k\|^2 \|g_{k+1}\|^2} \right] \\ &= \xi_{k+1}^2 \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \left[\frac{1}{1 + \sigma/4} - \frac{(g_k^T d_k)^2}{\|d_k\|^2 \|g_{k+1}\|^2} \right]. \end{aligned} \quad (۴۰.۲)$$

اما بر اساس لم ۲-۶ نتیجه می شد $\lim_{k \rightarrow \infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = 0$. حال اگر $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$ درست نباشد، آنگاه $\lim_{k \rightarrow \infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2 \|g_{k+1}\|^2} = 0$. بنابراین، از رابطه (۴۰.۲) نتیجه گرفته می شود که

$$\frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq c \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}.$$

واضح است که این موضوع به همراه لم ۲-۶ نشان می دهد نتیجه مورد نظر به دست آمده است. بنابراین، (۳۴.۲) برقرار است. $\square$

نتیجه ۲-۸. فرض کنید که

$$\sum_{k=1}^{\infty} \frac{1}{\|d_k\|^2} = \infty \quad (۴۱.۲)$$

تحت نتیجه لم ۲-۷ برقرار باشد. آنگاه، $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$.

اثبات. فرض کنید که ثابت ε برای هر $k \geq 1$ چنان وجود داشته باشد که $\|g_k\| \geq \varepsilon$ ، آنگاه لم ۲-۷ نتیجه می دهد که رابطه (۳۴.۲) برقرار است. سپس خواهیم داشت $\sum_{k=1}^{\infty} \frac{1}{\|d_k\|^2} \leq \frac{1}{\varepsilon^4} \sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$. با این حال، این موضوع با رابطه (۴۱.۲) در تناقض است پس باید رابطه (۳۳.۲) برقرار باشد. $\square$

قضیه ۲-۹. دنباله های $\{x_k\}$ و $\{d_k\}$ تولید شده توسط الگوریتم SH را در نظر بگیرید. فرض کنید فرض ۲-۴ برقرار باشد و عدد ثابت $0 < \varrho$ وجود داشته باشد به طوری که $\varrho \|g_{k+1}\| \|s_k\| \geq |g_{k+1}^T d_k|$. اگر f کوژ شدید باشد، آنگاه $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$.

اثبات. با توجه به کوژ شدید بودن تابع f ثابت $\mu > 0$ چنان وجود دارد که $[g(y) - g(x)]^T(y - x) \geq \mu\|y - x\|^2$, $\forall x, y \in \mathbb{R}^n$ با استفاده از پیوستگی $g(x)$ به دست می‌آید که $\|g_k\| \leq \gamma$. به‌طور مشابه، با اعمال رابطه (۲.۲) و فرض ۲-۴ نتیجه گرفته می‌شود که $\|y_k\| \leq \|s_k\|$ و $s_k = \alpha_k d_k$. بنابراین، رابطه $(g_{k+1} - g_k)^T(\alpha_k d_k) \geq \mu\|x_{k+1} - x_k\|^2 = \mu\alpha_k^2\|d_k\|^2$ برقرار است. در نهایت، از نامساوی $0 < \|s_k\|\|d_k\| \leq \mu\alpha_k\|d_k\|^2 = d_k^T y_k$ نتیجه می‌شود $\frac{L\gamma}{\mu\|d_k\|} \leq \left| \frac{g_{k+1}^T y_k}{d_k^T y_k} \right| = |\beta_k^{HS}|$. با ترکیب روابط (۲۵.۲) و (۲۶.۲) و شرط $\|y_k\| \leq L\|s_k\|$ برای ξ_{k+1} به دست می‌آید

$$|\xi_{k+1}| \leq 1 + 2 \frac{\|y_k\|}{\|s_k\|} \left| \frac{g_{k+1}^T s_k}{g_{k+1}^T y_k} \right| \leq 1 + 2L \frac{\|g_{k+1}\|\|s_k\|}{\varrho\|g_{k+1}\|\|s_k\|} = 1 + \frac{2L}{\varrho} := \tau.$$

بنابراین نتیجه می‌شود

$$\begin{aligned} \|d_{k+1}\| &= \|- \xi_{k+1} g_{k+1} + \beta_k^{HS} d_k\| \\ &\leq |\xi_{k+1}|\|g_{k+1}\| + |\beta_k^{HS}|\|d_k\| \\ &\leq \gamma\tau + \frac{L\gamma}{\mu}. \end{aligned}$$

این رابطه نشان می‌دهد $\sum_{k=1}^{\infty} \frac{1}{\|d_k\|^2} = \infty$. در نتیجه، اثبات با استفاده از نتیجه ۲-۸ کامل می‌شود. $\square$

تذکر ۱۰-۲. از آنجا که f کوژ شدید است و بر اساس تحلیل ارائه‌شده در [۴۸]، با استفاده از نامساوی کوشی شوارتز و این حقیقت که $d_k^T y_k \geq \mu\|s_k\|^2$ نامساوی $\|y_k\| \geq \mu\|s_k\|$ برقرار است. این رابطه نشان می‌دهد که اگر $|g_{k+1}^T y_k| = \|g_{k+1}\|\|y_k\|\cos \langle g_{k+1}, y_k \rangle| \geq \mu\|g_{k+1}\|\|s_k\|\cos \langle g_{k+1}, y_k \rangle|$ درواقع محدودیت زیادی بر روی پارامتر ϱ اعمال نمی‌کند. آنگاه، $|g_{k+1}^T y_k| \geq \varrho\|g_{k+1}\|\|s_k\|$.

۱.۳.۲ تحلیل همگرایی برای توابع عمومی

در این زیر بخش، برای فرمول‌بندی یک الگوریتم کلی برای حل مسائل با ابعاد بزرگ، فرض می‌شود یک روش گرادیان مزدوج طیفی برقرار است که تساوی موجود در رابطه (۲۱۰.۲) را برآورده می‌کند و یک تقریب B_{k+1} از هسیان حقیقی $\nabla^2 f(x_{k+1})$ بررسی می‌شود که از شرط سکانتی اصلاح‌شده ارائه‌شده در [۲۷، ۵۰] پیروی می‌کند. این شرط به صورت

$$B_{k+1} s_k = z_k, \quad z_k = y_k + h\|g_k\|^r s_k, \quad r > 0, h > 0 \quad (۴۲.۲)$$

تعریف می‌شود. با پیروی از مفهومی که در [۱، ۱۳، ۴۸] پیشنهاد شده‌است و استفاده از معادله سکانتی اصلاح‌شده در رابطه (۴۲.۲) جهت جستجو برای توابع عمومی غیرخطی به صورت

$$d_{k+1} = -\xi_{k+1}^{GSHS} g_{k+1} + \beta_k d_k, \quad \forall k \geq 0 \quad (43.2)$$

خواهد بود که در آن

$$\xi_{k+1}^{GSHS} = \max \left\{ 1, 1 + t \frac{s_k^T g_{k+1}}{g_{k+1}^T z_k} \right\}, \quad t = \frac{z_k^T s_k}{\|s_k\|^2} + \frac{\|z_k\|}{\|s_k\|} \quad (44.2)$$

و عبارت زیر با $\beta_k^{GHS} = \frac{g_{k+1}^T z_k}{d_k^T z_k}$

$$\beta_k = \begin{cases} \beta_k^{GHS}, & |g_{k+1}^T z_k| \leq \|g_{k+1}\|^2, \\ 0, & \text{در غیر این صورت.} \end{cases} \quad (45.2)$$

با این ساختار، شرط هم‌ارزی D-L نیز برآورده می‌شود. علاوه بر این، $\xi_{k+1} \geq 1$ برای هر $k \geq 0$ برقرار است. در ادامه، الگوریتم و تحلیل‌های نظری مربوط به روش HS طیفی اصلاح‌شده (GSHS) ارائه می‌شود.

الگوریتم ۲ GSHS

-
- گام ۱: $x_0 \in \mathbb{R}^n$ و پارامترهای $0 < \delta < \sigma < 1$ را انتخاب کن. مقداردهی اولیه به صورت $d_0 = -g_0$ و $\alpha_0 = 1$ انجام شود.
- گام ۲: الگوریتم را زمانی که $\|g_k\| \leq \epsilon$ متوقف کن.
- گام ۳: اندازه گام α_k را تعیین کن به طوری که شرایط جست‌وجوی ولف قوی (۳.۲) و (۵.۲) را برآورده کند.
- گام ۴: z_k را از (۴۲.۲) و پارامترهای ξ_{k+1}^{GSHS} و t را از β_k^{GHS} و (۴۴.۲) و (۴۵.۲) محاسبه کن.
- گام ۵: جهت جستجو را با $d = -\xi_{k+1}^{GSHS} g_{k+1} + \beta_k^{GHS} d_k$ محاسبه کن. شرط $|g_{k+1}^T g_k| > 0.2 \|g_{k+1}\|^2$ را بررسی کن. اگر برقرار بود، مقداردهی مجدد انجام بده، یعنی $d_{k+1} = -g_{k+1}$ در غیر این صورت، $d = d_{k+1}$.
- گام ۶: تکرار جدید را به صورت $x_{k+1} = x_k + \alpha_k d_k$ تولید کن.
- گام ۷: مقدار $k = k + 1$ را قرار بده و به گام ۲ برو.
-

لم ۱۱-۲. فرض کنید z_k براساس رابطه (۴۲.۲) تعریف شده باشد رابطه زیر همیشه برقرار است

$$\frac{|g_{k+1}^T d_k|}{|d_k^T z_k|} \leq \frac{\sigma}{1 - \sigma}. \quad (46.2)$$

اثبات. با استفاده از رابطه (۴۲.۲) و جایگذاری $s_k = \alpha_k d_k$ رابطه زیر حاصل می شود

$$d_k^T z_k = d_k^T y_k + h \|g_k\|^r \alpha_k \|d_k\|^2 \geq d_k^T y_k.$$

با ترکیب نامساوی فوق با روابط (۵.۲) و (۴.۲) خواهیم داشت

$$\begin{aligned} d_k^T z_k &\geq d_k^T y_k \\ &\geq -(1 - \sigma) d_k^T g_k \\ &\geq \frac{(1 - \sigma)}{\sigma} |g_{k+1}^T d_k|. \end{aligned}$$

این نتیجه می دهد $\frac{|g_{k+1}^T d_k|}{|d_k^T z_k|} \leq \frac{\sigma}{1 - \sigma}$ همان طور که خواسته شد برقرار است. $\square$

قضیه ۲-۱۲. فرض کنید $\bar{\xi}_{k+1}$ مطابق با رابطه (۴۴.۲) تعریف شده باشد و الگوریتم GSHS دنباله های $\{d_k\}$ و $\{g_k\}$ را تولید کند. آنگاه، ثابت $\bar{\rho} > 0$ وجود دارد به طوری که شرط زیر برقرار است

$$d_{k+1}^T g_{k+1} \leq -\bar{\rho} \|g_{k+1}\|^2, \quad \forall k \geq 0. \quad (۴۷.۲)$$

اثبات. با استفاده از تعریف d_{k+1} در رابطه (۴۳.۲) $g_{k+1}^T d_{k+1} = -\bar{\xi}_{k+1} \|g_{k+1}\|^2 + \frac{g_{k+1}^T z_k}{d_k^T z_k} g_{k+1}^T d_k$ (۴۳.۲) می شود. با استفاده از روابط (۴۵.۲) و (۴۶.۲) داریم

$$\begin{aligned} g_{k+1}^T d_{k+1} &\leq -\bar{\xi}_{k+1} \|g_{k+1}\|^2 + \frac{|g_{k+1}^T z_k|}{|d_k^T z_k|} |g_{k+1}^T d_k| \\ &\leq -(\bar{\xi}_{k+1} - \frac{\sigma}{1 - \sigma}) \|g_{k+1}\|^2. \end{aligned}$$

ویژگی نزول کافی با $\bar{\rho} = \bar{\xi}_{k+1} - \frac{\sigma}{1 - \sigma}$ و $\sigma \in (0, 0.4]$ برقرار می شود. این عبارت تضمین می کند که $\bar{\rho} > 0$ و نتیجه می دهد $\bar{\xi}_{k+1} \geq 1$. $\square$

برای دستیابی به همگرایی سراسری در توابع عمومی به لم هایی که از بخش های [۴۸، ۵۳] به دست آمده اند، نیاز می شود. لم ها عبارتند از،

لم ۲-۱۳. فرض کنید دنباله های $\{x_k\}$ ، $\{g_k\}$ و $\{d_k\}$ توسط الگوریتم GSHS تولید شده باشند. اگر فرض

۲-۴ برقرار باشد و ثابت $\varrho > 0$ چنان وجود داشته باشد که $|g_{k+1}^T z_k| \geq \bar{\varrho} \|g_{k+1}\| \|s_k\|$ برقرار باشد، آنگاه

$$\|d_{k+1}\| < G, \quad G > 0 \quad \forall k \geq 0. \quad (۴۸.۲)$$

تذکر ۲-۱۴. از آنجا که $|g_{k+1}^T z_k| = \|g_{k+1}\| \|s_k\| \cos \angle(g_{k+1}, z_k) < \|g_{k+1}\| \|s_k\|$ و $\|z_k\| \leq (L + h\gamma^r) \|s_k\|$ ، آنگاه انتخاب پارامتر $\bar{\rho}$ محدودیت زیادی ایجاد نمی‌کند.

قضیه ۲-۱۵. فرض کنید الگوریتم GSHS دنباله‌های $\{x_k\}$ و $\{d_k\}$ را تولید کند به طوری که جهت جست‌وجوی d_k نزولی باشد و α_k شرط ولف قوی را برآورده کند در این صورت، فقط یکی از دو عبارت زیر برقرار است

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0, \quad (۴۹.۲)$$

$$\sum_{k=1}^{\infty} \frac{\|g_{k+1}\|^4}{\|d_k\|^4} < +\infty \text{ و}$$

اثبات. فرض کنید رابطه (۴۹.۲) برقرار نباشد، بدین معنی که g_k از صفر دور باشد. یعنی یک ثابت $\bar{\varepsilon} > 0$ وجود دارد به طوری که

$$\|g_k\| \geq \bar{\varepsilon}, \quad \forall k \geq 1. \quad (۵۰.۲)$$

از رابطه (۴۲.۲) و فرض ۲-۴ داریم

$$\|z_k\| = L\|s_k\| + h\|g_k\|^r \|s_k\| \leq (L + h\gamma^r) \|s_k\|. \quad (۵۱.۲)$$

سپس، با استفاده از نامساوی کوشی-شوارتز، روابط (۴۳.۲) و (۴۴.۲) همراه با (۵۱.۲) خواهیم داشت

$$|t| \leq \frac{|z_k^T s_k|}{\|s_k\|^2} + \frac{\|z_k\|}{\|s_k\|} \leq 2 \frac{\|z_k\|}{\|s_k\|} \leq 2(L + h\gamma^r) := M$$

حال، $|\xi_{k+1}| \leq 1 + M \left| \frac{g_{k+1}^T s_k}{g_{k+1}^T z_k} \right| \leq 1 + M \frac{\|g_{k+1}\| \|s_k\|}{\bar{\varrho} \|g_{k+1}\| \|s_k\|} = 1 + \frac{M}{\bar{\varrho}} := \bar{\tau}$ ، روابط (۴۰.۲) و (۴۷.۲) می‌توان نوشت

$$d_k^T y_k \geq -(1 - \sigma) d_k^T g_k \geq \bar{\rho}(1 - \sigma) \|g_k\|^2 > 0. \quad (۵۲.۲)$$

بنابراین، با استفاده از روابط (۵۱.۲) و (۵۲.۲) و این حقیقت که $\|s_k\| = \|x_{k+1} - x_k\| \leq 2B$ داریم

$$|\beta_k^{GHS}| = \left| \frac{g_{k+1}^T z_k}{d_k^T z_k} \right| \leq \frac{\|g_{k+1}\| \|z_k\|}{|d_k^T y_k|} \leq \frac{\gamma(L + h\gamma^r) \|s_k\|}{\bar{\rho}(1 - \sigma) \|g_k\|^2} \leq \frac{2\gamma(L + h\gamma^r)B}{\bar{\rho}(1 - \sigma)\bar{\varepsilon}} := \bar{M}$$

$\alpha_k \geq \vartheta > 0$ نتیجه می‌شود $d_{k+1} = -\xi_{k+1}g_{k+1} + \beta_k \frac{s_k}{\alpha_k}$ ، یعنی

$$\begin{aligned} \|d_{k+1}\| &\leq |\xi_{k+1}| \|g_{k+1}\| + |\beta_k| \frac{\|s_k\|}{\alpha_k} \\ &\leq \bar{\tau}\gamma + \bar{M} \frac{\gamma B}{\alpha_k} \leq \bar{\tau}\gamma + \bar{M} \frac{\gamma B}{\vartheta}. \end{aligned}$$

این نشان می‌دهد که ادعای ارائه‌شده در رابطه (۴۸.۲) برقرار است. درعین حال، از روابط (۴۷.۲)، (۴۸.۲) و (۵۰.۲) به دست می‌آید، $(d_{k+1}^T g_{k+1})^2 \geq \bar{\rho}^2 \|g_{k+1}\|^4$ ، با تقسیم بر $\|d_{k+1}\|^2$ و جمع کردن نتیجه می‌شود $\bar{\rho}^2 \sum_{k=0}^{\infty} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \geq \sum_{k=0}^{\infty} \frac{\bar{\rho}^2 \varepsilon^4}{G^2} \geq \frac{\bar{\rho}^2 \varepsilon^4}{G^2} \sum_{k=0}^{\infty} 1 = +\infty$ □

۴.۲ نتایج عددی

این بخش قدرت روش GSHS را در مقایسه با برخی روش‌های گرادیان مزدوج گزارش می‌دهد. مقداردهی‌های عددی بر اساس تابع‌های آزمایش با ابعادی بین ۲ تا ۵۰۰۰۰۰ انجام شده‌است، همان طور که در [۳، ۳۲] مورد بررسی قرار گرفته‌است. برای تحلیل نتایج، الگوریتم‌های زیر بررسی می‌شوند،

• الگوریتم GSHS پیشنهادشده با $h = 10^{-3}$ و r محاسبه‌شده به صورت زیر است

$$r = \begin{cases} 1, & \|g_k\| \geq 1, \\ 3, & \text{درغیراین صورت.} \end{cases}$$

• الگوریتم TSCG فرامرزی و امینی [۱۴] جهت جستجو را $d_k = -\theta_k g_k + \beta_k^{MHS} d_{k-1} - u_k z_{k-1}$ به‌کار برده است که در آن $\beta_k^{MHS} = \frac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}}$ ، $u_k = \frac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}}$ و θ_k با استفاده از روابط (۱۳.۲) و (۱۴.۲) با $\eta_1 = 0.001$ ، $\eta_2 = 10$ ، $z_{k-1} = y_{k-1} + 0.001 \|g_{k-1}\|^r s_{k-1}$ ، $\delta = 0.01$ و $\sigma = 0.1$ از روابط (۳.۲) و (۵.۲) مشخص شده‌است. r نیز از [۵۲] گرفته شده‌است.

• الگوریتم 1 NDTTHSCG ونگ [۴۸] جهت جستجو را $d_{k+1} = -\theta_k^{HS} g_{k+1} + \beta_k^{HS} d_k - \frac{g_{k+1}^T d_k}{y_k^T d_k} y_k$ تعریف کرده‌است که در آن θ_{k+1}^{HS} از بخش [۴۸] گرفته شده‌است، β_k^{HS} از (۷.۲) با مقادیر $\delta = 0.0001$ و $\sigma = 0.80$ در روابط (۳.۲) و (۵.۲) تعریف شده‌است.

• الگوریتم 2 NDTTHSCG ونگ [۴۸] جهت جستجو رابه صورت زیر تعریف کرده‌است که در آن θ_{k+1}^{NHS} و β_k^{NHS} از بخش [۴۸] به ترتیب با مقادیر $\delta = 0.0001$ و $\sigma = 0.80$ از روابط (۳.۲) و (۵.۲) گرفته

شده است

$$d_{k+1} = -\theta_k^{NHS} g_{k+1} + \beta_k^{NHS} d_k - \frac{g_{k+1}^T d_k}{z_k^T d_k} y_k.$$

• کاهش- CG ، الگوریتم هاگر و ژانگ در بخش [۱۹] که در آن d_k از رابطه (۶.۲) به دست می آید و β_k مطابق با تعریف ارائه شده در [۱۹] با مقادیر $\delta = 0.001$ و $\sigma = 0.5$ در روابط (۳.۲) و (۵.۲) تعیین می شود.

• الگوریتم MSTCG امینی و فرامیزی در بخش [۱] جهت جستجو را به صورت زیر تعریف کرده است که در آن $w_k = \max\{\psi \|d_{k-1}\| \|y_{k-1}\|, d_{k-1}^T y_{k-1}, \|g_{k-1}\|^2\}$ با $\psi = 0.2$ و $\delta = 0.01$ ، $\sigma = 0.1$ از روابط (۳.۲) و (۵.۲)

$$d_{k+1} = -\theta_k g_k + \frac{g_k^T y_{k-1}}{w_k} d_{k-1} - \frac{g_k^T d_{k-1}}{w_k} y_{k-1}.$$

• الگوریتم MF طیفی از حاتم و همکاران در بخش [۳۳] جهت جستجو را $d_{k+1} = -g_{k+1} + \beta_k d_k$ تعریف کرده است که β_k مطابق با [۳۳] تعیین می شود و $\delta = 0.01$ و $\sigma = 0.1$ از روابط (۳.۲) و (۵.۲) به کار گرفته می شوند.

تمامی کدها در MATLAB R2022a پیاده سازی شده است و روی یک لپ تاپ Dell Intel (R) Core i7-1195G7 با ۱۶ گیگابایت رم و پردازنده ۹۰.۲ گیگاهرتز اجرا شده اند. برای روش GSHS $\delta = 10^{-3}$ و $\sigma = 0.4$ از روابط (۳.۲) و (۵.۲) تنظیم شده اند و شرط توقف برای تمامی روش ها $\|g_k\| \leq 10^{-6}$ در نظر گرفته شده است. علامت *** نشان دهنده حالتی است که در آن شرط توقف برآورده نشده است. عملکرد محاسباتی به طور دقیق در جداول ۱-۴ ارائه شده است. برای تفسیر و تحلیل عملکرد روش هایی که توصیف شده اند از تکنیک پروفایل عملکرد ارائه شده توسط دولن و موری در بخش [۱۲] استفاده می شود.

فرض کنید P مجموعه ای از n_p مسئله آزمون و S مجموعه ای از n_s حل کننده مورد استفاده در مقایسه باشد. بنابراین، پروفایل عملکرد معیاری برای ارزیابی عملکرد در مسئله $p \in P$ و حل کننده $s \in S$ است که نشان دهنده تعداد ارزیابی های تابع (FE)، تعداد تکرارها (NI) یا زمان سپری شده (ET) نسبت به سایر حل کننده های $s \in S$ روی مجموعه ای از مسائل $p \in P$ می باشد. آنگاه، معیار نسبت عملکرد برای مقایسه و ارزیابی کارایی به صورت زیر تعریف می شود

$$r_{p,s} = \frac{f_{p,s}}{\min\{f_{p,s} : s \in S, p \in P\}}.$$

S/N	Test functions	Dimm	Initial point	GSHS			CG _D ESCENT			MSTCG			NDTHSCG ^١			NDTHSCG ^٢			MF			TSCG		
				NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU
P1	EXT PENALTY	100	(١١, ..., ١١)	٩	٤٧	٠.٠٠٠٥	٨	٢٥	٠.٠٠٠٥	٩	٢٨	٠.٠٠٠٦	٩	٢٨	٠.٠٠٠٦	٩	٢٨	٠.٠٠٠٦	٩	٢٨	٠.٠٠٠٦	٥	٣٥	٠.٠٠٠٦
P2	EXT PENALTY	٥٠٠	(١١, ..., ١١)	٧	٣٨	٠.٠٠٠٣	٧	٣٩	٠.٠٠٠٦	٧	٣١	٠.٠٠٠٧	٧	٣٨	٠.٠٠٠٧	٧	٣٨	٠.٠٠٠٧	٧	٣٨	٠.٠٠٠٧	٤	٣٠	٠.٠٠٠٦
P3	EXT PENALTY	1000	(١١, ..., ١١)	٧	٣٨	٠.٠٠٠٣	٧	٣٩	٠.٠٠٠٦	٥	٢٦	٠.٠٠٠٨	٧	٣٨	٠.٠٠٠٩	٧	٣٨	٠.٠٠٠٩	٧	٣٨	٠.٠٠٠٩	٥	٥٤	٠.٠٠٠٥
P4	EXT PENALTY	٢٥٠٠	(١١, ..., ١١)	٦	٣٤	٠.٠٠١٥	٧	٣٩	٠.٠٠١٥	٦	٢٢	٠.٠٠١٥	***	***	***	***	***	***	٨	٤٦	٠.٠٠١٥	٤	٣٠	٠.٠٠١٧
P5	EXT PENALTY	٥٠٠٠	(١١, ..., ١١)	٦	٣٤	٠.٠٠٠٣	***	***	***	٥	٢٢	٠.٠٠٢٣	٨	٤٦	٠.٠٠٣٧	٨	٤٦	٠.٠٠٣٧	٨	٤٦	٠.٠٠٣٧	٤	٢٩	٠.٠٠٣٢
P6	EXT PENALTY	10000	(١١, ..., ١١)	٦	٣٤	٠.٠٠٥٩	***	***	***	٥	٢٢	٠.٠٠٤٧	٨	٤٦	٠.٠٠٥٩	٨	٤٦	٠.٠٠٥٩	٨	٤٦	٠.٠٠٥٩	٥	٢٤	٠.٠٠٦٢
P7	DIAGONAL Δ	10	(1, ..., 1)	٢	1٣	٠.٠٠٢٧	٢	1٣	٠.٠٠٩٧	٢	11	٠.٠٠٤٧	٢	11	٠.٠٠٧٦	٢	11	٠.٠٠٧٦	٢	11	٠.٠٠٧٦	٢	11	٠.٠٠٦٩
P8	DIAGONAL Δ	٥٠٠٠	(1, ..., 1)	٢	1٣	٠.٠٠٩	٢	1٣	٠.٠١٣٦	٢	11	٠.٠١٩٤	٢	11	٠.٠٢٦٥	٢	11	٠.٠٢٦٥	٢	11	٠.٠٢٦٥	٢	11	٠.٠٢٣١
P9	DIAGONAL Δ	10000	(1, ..., 1)	٢	1٣	٠.1٧٧٧	٢	1٣	٠.1٩٤٨	٢	11	٠.1٥٤٩	٢	11	٠.1٨11	٢	11	٠.1٩٤٨	٢	11	٠.1٩٤٨	٢	11	٠.1٨٠٩
P10	TRECANNI	٢	(٥, ٥)	٦	٣٣	٠.٠٠٠٤	11	٥٢	٠.٠٠٠٣	1٢	٧٥	٠.٠٠٠٩	٩	٤٣	٠.٠٠٠٨	٩	٤٣	٠.٠٠٠٨	1٢	٧٣	٠.٠٠٠٩	٩	٤٧	٠.٠٠٠٨
P11	TRECANNI	٢	(٥, ٥)	10	٦٠	٠.٠٠٠٣	11	٦٨	٠.٠٠٠٣	11	٧٢	٠.٠٠٠٣	10	٦1	٠.٠٠٠٣	10	٦1	٠.٠٠٠٣	1٢	10٨	٠.٠٠٠٣	10	٦٥	٠.٠٠٠٣
P12	Q. PENALTY ١	1000	(٢, ..., ٢)	٩	٤٦	٠.٠٠٢1	1٥	٩٢	٠.٠٠1٧	***	***	***	1٦	٩1	٠.٠٠٢	1٦	٩1	٠.٠٠٢٦	***	***	***	1٦	٩1	٠.٠٠٢٦
P13	Q. PENALTY ١	1٢٠٠	(٢, ..., ٢)	٨	٤٢	٠.٠٠٢	11	٦٤	٠.٠٠1٧	1٧	٩٧	٠.٠٠٢٨	1٤	٨٥	٠.٠٠٢٨	1٤	٨٥	٠.٠٠٢٨	10	٥٠	٠.٠٠٢٨	***	***	***
P14	Q. PENALTY ١	1٣٠٠	(٢, ..., ٢)	11	٥٤	٠.٠٠٢٥	1٥	٩٠	٠.٠٠٢٩	1٧	11٦	٠.٠٠٣٨	10	٦٤	٠.٠٠1٧	10	٦٤	٠.٠٠1٧	1٢	٥٨	٠.٠٠١٦	1٢	11٨	٠.٠٠٣٢
P15	Q. PENALTY ١	1٥٠٠	(٢, ..., ٢)	10	٥٠	٠.٠١٢٥	1٦	٩٥	٠.٠٠٣٦	10	٥٥	٠.٠٠٢٤	1٤	٨٤	٠.٠٠٣٦	1٤	٨٤	٠.٠٠٣٦	***	***	***	1٤	٧٨	٠.٠٠٢٤
P16	Q. FUNCTION ١	٢	(٢, ..., ٢)	٤	1٧	٠.٠٠٠٩	٢	٩	٠.٠٠٠٩	٢	٩	٠.٠٠٠٩	٢	٩	٠.٠٠٠٩	٢	٩	٠.٠٠٠٩	1٢	٥٣	٠.٠٠٠٩	٣	1٣	٠.٠٠٠٩
P17	Q. FUNCTION ١	٤	(٢, ..., ٢)	٢٠	٩٤	٠.٠٠1٧	٢٨	1٢٣	٠.٠٠1٩	1٥	٩٧	٠.٠٠٢٥	1٧	1٣٢	٠.٠٠1٥	1٧	1٣٢	٠.٠٠1٥	1٢	٩٤	٠.٠٠1٥	٣٠	1٢٤	٠.٠٠1٩
P18	Q. FUNCTION ١	٦	(٢, ..., ٢)	1٤	٦٧	٠.٠٠٠٧	٣٠	111	٠.٠٠1٣	1٧	1٣1	٠.٠٠1٦	1٤	10٨	٠.٠٠1٣	1٤	10٨	٠.٠٠1٣	1٢	٦٧	٠.٠٠٠٧	1٢	٩٠	٠.٠٠٢٩
P19	Q. FUNCTION ١	٨	(٢, ..., ٢)	1٢	٦٧	٠.٠٠١٦	٢٢	٩٥	٠.٠٠٢٥	1٦	1٣1	٠.٠٠1٦	1٤	1٥٥	٠.٠٠1٦	1٤	1٥٥	٠.٠٠1٦	1٢	1٢٣	٠.٠٠1٦	1٢	1٨1٧	٠.٠٠1٦
P20	Q. FUNCTION ١	1٢	(٢, ..., ٢)	1٢	٦1	٠.٠٠٢٥	***	***	***	1٥	1٢٦	٠.٠٠1٦	***	***	***	***	***	***	1٩	1٣٥	٠.٠٠1٩	***	***	***
P21	Q. FUNCTION ٢	10	(1, ..., 1)	٣	1٣	٠.٠٠٣٨	٤	1٧	٠.٠٠1٥	٢	10	٠.٠٠٢٥	٣	1٣	٠.٠٠٣٦	٣	1٣	٠.٠٠٣٦	٣	1٣	٠.٠٠٣٦	٢	٩	٠.٠٠٣٦
P22	Q. FUNCTION ٢	٥٠٠	(1, ..., 1)	٣	1٣	٠.٠٠٢1	٤	1٧	٠.٠٠٣1	٢	٢٢	٠.٠٠٣1	٣	1٣	٠.٠٠٢٤	٣	1٣	٠.٠٠٢٤	٣	1٣	٠.٠٠٢٤	٢	٩	٠.٠٠٢٧
P23	Q. FUNCTION ٢	1000	(1, ..., 1)	٣	1٣	٠.٠٠٥٢	٤	1٧	٠.٠٠٣٨	1	٢	٠.٠٠٣٨	٣	1٣	٠.٠٠٣٦	٣	1٣	٠.٠٠٣٦	٣	1٣	٠.٠٠٣٦	٢	1٣	٠.٠٠٣٦
P24	Q. FUNCTION ٢	٥٠٠٠	(1, ..., 1)	٣	1٣	٠.٠٠٤٩	٤	1٧	٠.٠٠٧٦	1	٢	٠.٠٠٥٨	٣	1٣	٠.٠٠٥٩	٣	1٣	٠.٠٠٧٨	٣	1٣	٠.٠٠٧٨	٢	1٣	٠.٠٠٨٢
P25	POWER	٢	(٥, ٥)	1٢	٤٩	٠.٠٠٠٣	1٥	٦1	٠.٠٠٠٨	٢	11	٠.٠٠٠٣	1٣	٥٣	٠.٠٠٠٣	1٣	٥٣	٠.٠٠٠٣	1٢	٤٩	٠.٠٠٠٣	٢	1٤	٠.٠٠٠٣
P26	POWER	٢	(٢, ٢)	1٣	٥٣	٠.٠٠٠٣	1٦	٦٥	٠.٠٠٠٣	٢	11	٠.٠٠٠٣	1٤	٥٧	٠.٠٠٠٣	1٤	٥٧	٠.٠٠٠٣	1٣	٥٣	٠.٠٠٠٣	٢	1٥	٠.٠٠٠٣
P27	ZETTL	٢	(1٣, ..., ١٣)	1٥	٥٩	٠.٠٠٠٣	1٣	11٨	٠.٠٠٠٣	٢٨	1٥٣	٠.٠٠٠٣	1٨	١٠٦	٠.٠٠٠٣	1٨	١٠٦	٠.٠٠٠٣	1٥	٥٩	٠.٠٠٠٣	٢	1٥	٠.٠٠٠٣
P28	TEST	٣	(1, 1, 1)	10	٨٥	٠.٠٠٠٦	10	٦٧	٠.٠٠٠٣	٢٠	1٣1	٠.٠٠٠٦	1٥	1٠٦	٠.٠٠٠٦	1٥	1٠٦	٠.٠٠٠٦	***	***	***	٢٥	1٣٨	٠.٠٠٢٨
P29	TEST	٣	(٢, ٢, ٢)	٧	٧٣	٠.٠٠٠٤	٢٩	1٥١	٠.٠٠١٩	1٢٦	1٥١	٠.٠٠٠٦	1٥	1٠٦	٠.٠٠٠٦	1٥	1٠٦	٠.٠٠٠٦	٧	٧٣	٠.٠٠٠٣	٢٩	1٦٧	٠.٠٠1٢
P30	TEST	٣	(٣, ٣, ٣)	1٣	1٥٩	٠.٠٠٠٩	٢٤	1٥٥	٠.٠٠٨1	1٤٤	١٥٩	٠.٠٠٢٣	1٥٥	1٥1	٠.٠٠٢٣	1٥٥	1٥1	٠.٠٠٢٣	1٣	1٥٩	٠.٠٠٢٣	1٣	1٥٩	٠.٠٠٢٣
P31	TEST	٣	(٥, ٥, ٥)	11	٩٣	٠.٠٠٥٧	٢٢	1٢٤	٠.٠٠1٢	1٢	1٨٨	٠.٠٠٢٥	1٥	1٥٣	٠.٠٠1٢	1٥	1٥٣	٠.٠٠1٢	1٢	٩٣	٠.٠٠1٢	1٢	1٨٨	٠.٠٠1٢
P32	SUM SQUARE	100	(٥, 1, ..., ٥, 1)	1٦	1٥٧	٠.٠٠٠٣	***	***	***	***	***	***	1٦	1٥٧	٠.٠٠٠٣	1٦	1٥٧	٠.٠٠٠٣	***	***	***	1٦	1٥٧	٠.٠٠٠٣
P33	SUM SQUARE	٥٠٠	(٥, 1, ..., ٥, 1)	10	111	٠.٠٠٠٣	***	***	***	***	***	***	10	111	٠.٠٠٠٣	10	111	٠.٠٠٠٣	***	***	***	10	111	٠.٠٠٠٣
P34	SUM SQUARE	1000	(٥, 1, ..., ٥, 1)	11	1٥٤	٠.٠٠٠٣	***	***	***	***	***	***	11	1٥٤	٠.٠٠٠٣	11	1٥٤	٠.٠٠٠٣	***	***	***	11	1٥٤	٠.٠٠٠٣
P35	SUM SQUARE	10000	(٥, 1, ..., ٥, 1)	1٢	1٥1	٠.٠٠٠٣	***	***	***	***	***	***	1٢	1٥1	٠.٠٠٠٣	1٢	1٥1	٠.٠٠٠٣	***	***	***	1٢	1٥1	٠.٠٠٠٣
P36	SUM SQUARE	٥٠٠٠	(٥, 1, ..., ٥, 1)	1٢	٩٥٣	٠.٠٠٠٣	***	***	***	***	***	***	1٢	٩٥٣	٠.٠٠٠٣	1٢	٩٥٣	٠.٠٠٠٣	***	***	***	1٢	٩٥٣	٠.٠٠٠٣

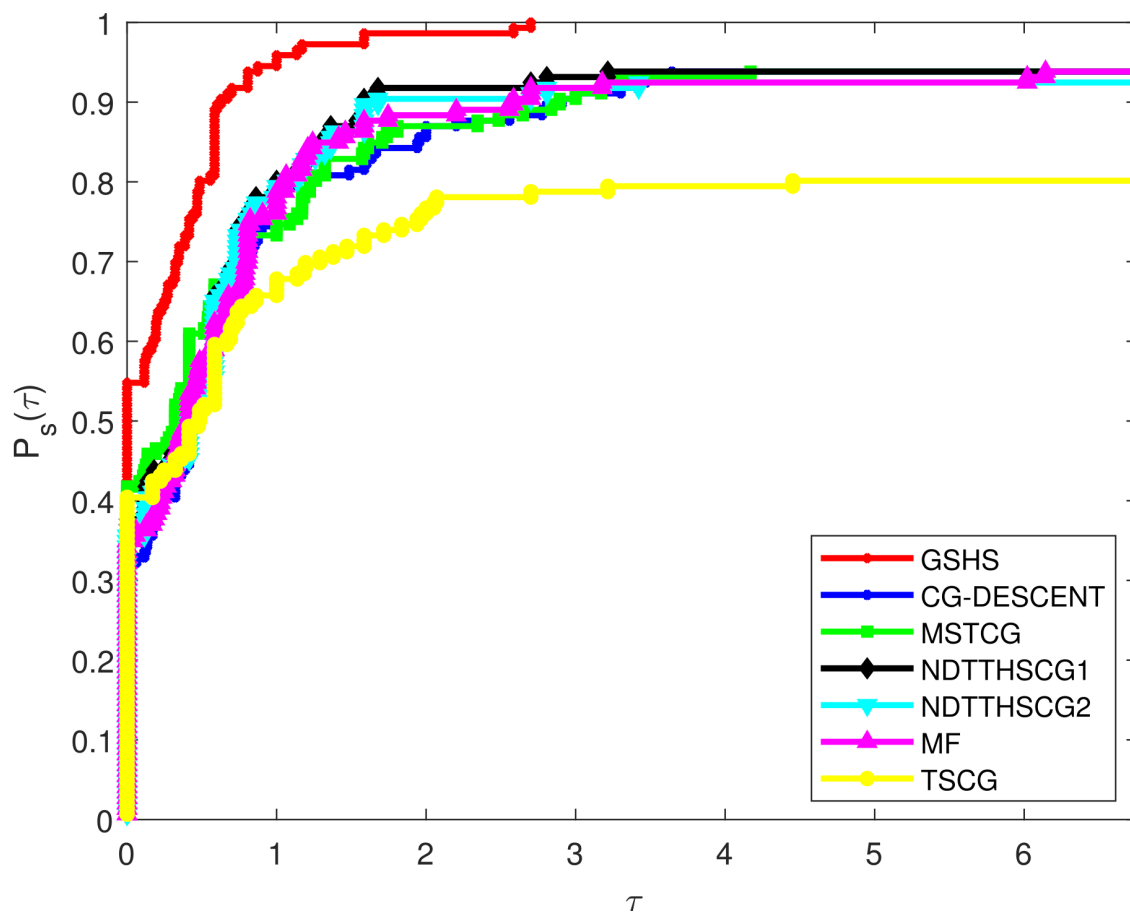
جدول ١.٢: نتایج عددی الگوریتمها

S/N	Test functions	Dimm	Initial point	GSHS			CG _D ESCENT			MSTCG			NDTHSCG ^١			NDTHSCG ^٢			MF			TSCG		
				NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU
P٣٧	DQUARTC	1٠٠	(٥/١٥, ..., ٥/١٥)	٢	1٤	٠.٠٠٢٨	٢	1٣	٠.٠١	٣	٢٥	٠.٠٠٧٤	٣	٢٣	٠.٠٠٥	٣	٢٣	٠.٠٠٥٦	٣	1٤	٠.٠٠٢٩	٣	٢٥	٠.٠٠٦1
P٣٨	DQUARTC	1٠٠٠٠	(٥/١٥, ..., ٥/١٥)	٣	٢٥	٠.٠11٢	٣	٢٣	٠.٠11٥	٤	٢٥	٠.٠٢11	٣	٢٣	٠.٠1٥٧	٣	٢٣	٠.٠1٥٧	٣	1٤	٠.٠1٢٨	٣	٢٥	٠.٠1٨٣
P٣٩	DQUARTC	٥٠٠٠	(٥/١٥, ..., ٥/١٥)	٣	٢٥	٠.٠٤٥٧	٣	٢٣	٠.٠٦٣٦	٤	٢٥	٠.٠٥٧٩	٣	٢٥	٠.٠٤٣٤	٣	٢٥	٠.٠٤٣٤	٣	٢٥	٠.٠٤٧٨	٣	٢٥	٠.٠٣٥٩
P٣٠	DQUARTC	1٠٠٠٠	(٥/١٥, ..., ٥/١٥)	٣	٢٥	٠.٠٩٩٩	٣	٢٣	٠.٠٩٧٥	٤	٢٥	٠.1٥٣٣	٣	٢٥	٠.٠٧٣٣	٣	٢٥	٠.٠٧٣٣	٣	٢٥	٠.٠٨٨٧	٣	٢٥	٠.٠٥٦٤
P٣1	DQUARTC	٥٠٠٠٠	(٥/١٥, ..., ٥/١٥)	٣	٢٥	٠.٢1٦٤	٣	٢٣	٠.٢1٦٤	٤	٢٥	٠.٥٥1٤	٣	٢٥	٠.٢٣٣	٣	٢٥	٠.٢1٩٦	٣	٢٥	٠.٢١٥٣	٣	٢٥	٠.٢٨٢٦
P٣٢	DQUARTC	1٠٠٠٠٠	(٥/١٥, ..., ٥/١٥)	٣	٢٥	٠.٦٣1٨	٣	٢٣	٠.٦٣٦٣	٤	٢٥	1.٠٨٢٦	٣	٢٥	٠.٦٣٦٣	٣	٢٥	٠.٦٣٦٣	٣	٢٥	٠.٨٨٧٣	٣	٢٥	٠.٥1٨1
P٣٣	MATYAS	٢	(1, 1, 1)	1	1٠	٠.٠٠٠٥	1	1٠	٠.٠٠٠٨	1	1٠	٠.٠٠1٣	1	1٠	٠.٠٠٠٣	1	1٠	٠.٠٠٠٥	1	1٠	٠.٠٠٠٣	1	1٠	٠.٠٠٠٣
P٣٤	MATYAS	٢	(٥, ٥)	1	1٠	٠.٠٠٠٣	1	1٠	٠.٠٠٠٥	1	1٠	٠.٠٠1٥	1	1٠	٠.٠٠٠٤	1	1٠	٠.٠٠٠٥	1	1٠	٠.٠٠٠٤	1	1٥	٠.٠٠٠٩
P٣٥	DIAGONAL ١	٢	(٢, ٢)	٥	٢٢	٠.٠٠1٣	1٤	٥٢	٠.٠٠٠٧	1٢	٥٠	٠.٠٠٠٩	11	٤٥	٠.٠٠1٨	11	٤٥	٠.٠٠٠٨	٢٢	٩٢	٠.٠١٥٣	1٣	٤٢	٠.٠٥٧1
P٣٦	DIAGONAL ١	٢	(-٢, -٢)	٥	٢٢	٠.٠٠1٣	٥	٥٢	٠.٠٠٠٧	٩	٢٩	٠.٠٠٠٥	٩	1٩	٠.٠٠٢٧	٩	1٩	٠.٠٠٢٧	11	٩٢	٠.٠٠٢٧	٩	1٩	٠.٠٠٢٧
P٣٧	DIAGONAL ١	٢	(-٥٢, -٥٢)	٩	1٨	٠.٠٠1٧	٥	٢1	٠.٠٠٣1	٧	٢٢	٠.٠٠٢٢	٤	1٧	٠.٠٠1٨	٩	1٧	٠.٠٠٢٤	٧	٢٩	٠.٠٠٢٣	٤	1٨	٠.٠٠٢٤
P٣٨	DIAGONAL ١	٢	(٥٢, ٥٢)	٩	1٥	٠.٠٠1٥	٥	٢1	٠.٠٠٢٨	٧	٢٩	٠.٠٠٢٥	٥	1٧	٠.٠٠٢٣	٩	1٧	٠.٠٠٢٤	٧	٢٩	٠.٠٠٢٥	٤	1٨	٠.٠٠٢٣
P٣٩	HAGER	1٠	(٢, ..., ٢)	1٧	٧٠	٠.٠٠٠٩	1٣	٤٢	٠.٠٠٠٩	1٩	1٠٧	٠.٠٠٠٩	1٩	٩٧	٠.٠٠٠٩	1٩	٩٧	٠.٠٠٠٩	٢٧	11٠	٠.٠٠٠٢	٢1	٩٨	٠.٠٠٠٤
P٤٠	HAGER	١٠	(٢, ..., ٢)	1٠	٢٤	٠.٠٠٠٩	1٠	٤٢	٠.٠٠٠٥	11	٢٨	٠.٠٠٠٩	1٠	٤٢	٠.٠٠٠٨	1٠	٤٢	٠.٠٠٠٥	1٧	٢1٠	٠.٠٠٠٥	1٠	٢٤	٠.٠٠٠٨
P٤1	HAGER	1٠	(٢, ..., ٢)	1٧	٧٠	٠.٠٠٠٤	1٣	٤٢	٠.٠٠٠٢	1٩	1٠٧	٠.٠٠٠٩	1٩	٩٧	٠.٠٠٠٩	1٩	٩٧	٠.٠٠٠٢	٢٧	11٠	٠.٠٠٠٤	٢1	٩٨	٠.٠٠٠٥
P٤٢	ZIRILLI	٢	(1, 1, 1)	٨	٢٨	٠.٠٠1٩	٩	٤1	٠.٠٠1	٩	٥٥	٠.٠٠1٨	٧	٢٣	٠.٠٠٠٢	٧	٢٣	٠.٠٠٠٢	٩	1٢1	٠.٠٠٠٩	1٠	٥1	٠.٠٠٠٢
P٤٣	ZIRILLI	٢	(٢, ٢)	٢	1٣	٠.٠٠٠٢	٩	٤٢	٠.٠٠٠٢	٧	٢٣	٠.٠٠٠٧	٥	٢1	٠.٠٠٠٢	٥	٢1	٠.٠٠٠٢	٩	1٢	٠.٠٠٠٩	٥	٢1	٠.٠٠٠٨
P٤٤	ZIRILLI	٢	(٢, ٢)	٥	٢٢	٠.٠٠٠٢	٧	٢1	٠.٠٠1٢	11	٥٢	٠.٠٠٠٩	٩	٢٧	٠.٠٠٠٢	٩	٢٧	٠.٠٠٠٢	٩	1٢٠	٠.٠٠٠٧	٧	٢٢	٠.٠٠٠٢
P٤٥	ZIRILLI	٢	(-٢, -٢)	٢	1٣	٠.٠٠٠٢	٥	٢1	٠.٠٠٠٢	٩	٥٢	٠.٠٠٠٨	٥	٢1	٠.٠٠٠٢	٥	٢1	٠.٠٠٠٢	٩	٢1	٠.٠٠٠1	٩	٢٤	٠.٠٠٠٢
P٤٦	RAYDAN ١	٢	(1, ..., 1)	٩	٢٨	٠.٠٠٠٣	٩	٢٩	٠.٠٠٥٢	٥	٥٢	٠.٠٠٠٨	٩	٢٩	٠.٠٠٥٨	٩	٢٩	٠.٠٠٠٩	٩	٢٨	٠.٠1٥1	٢٠	٥٥	٠.٠٠٥٥
P٤٧	RAYDAN ١	٢	(1, ..., 1)	1٢	٢٨	٠.٠٠٥٥	11	٢٢	٠.٠٠٥٨	11	٧٢	٠.٠٠٥٩	11	٧٢	٠.٠٠٥٧	11	٧٢	٠.٠٠٥٩	٢٢	1٢٩	٠.٠1٢٩	11	٧٠	٠.٠٠٥٦
P٤٨	RAYDAN ١	1٠	(1, ..., 1)	٢٥	٩٨	٠.٠٠1	1٠	111	٠.٠1٠٢	1٨	٩1	٠.٠٠1٢	1٧	٩٥	٠.٠٠1٧	1٧	٩٥	٠.٠1٠1	٥٧	٢1٠	٠.٠٢٩٩	1٧	٩٥	٠.٠٠٧٦
P٤٩	RAYDAN ٢	1٠٠٠٠	(٥/٢, ..., ٥/٢)	٢	٩	٠.٠٠٠1	٢	٩	٠.٠11٧٧	٢	1٠	٠.٠1٢٢	٢	٩	٠.٠11٢٧	٢	٩	٠.٠11٥	٢	٩	٠.٠1٢٥	٢	1٠	٠.٠1٢٤
P٥٠	RAYDAN ٢	٥٠٠٠	(٥/٢, ..., ٥/٢)	٢	٩	٠.٠٢٥	٢	٩	٠.٠٢٣٢	٢	1٠	٠.٠٢٠1	٢	٩	٠.٠٢٠٢	٢	٩	٠.٠٢٣٢	٢	٩	٠.٠٢٥٢	٢	1٠	٠.٠٢٥٩
P٥1	RAYDAN ٢	1٠٠٠٠	(٥/٢, ..., ٥/٢)	٢	٩	٠.٠٥٩٧	٢	٩	٠.٠٧	٢	1٠	٠.٠٩٨٥	٢	٩	٠.٠٥٢٢	٢	٩	٠.٠٥٢٢	٢	٩	٠.٠٧٥	٢	1٠	٠.٠٥٢٥
P٥٢	FLETCHCR	1٠	(٠.٩٩, ..., ٠.٩٩)	٢	11	٠.٠٠٠٢	٢	11	٠.٠٠٠٢	٢	11	٠.٠٠٠٧	٢	11	٠.٠٠٠٢	٢	11	٠.٠٠٠٥	٢	11	٠.٠٠٠٥	٢	11	٠.٠٠٠٤
P٥٣	FLETCHCR	1٠٠	(٠.٩٩, ..., ٠.٩٩)	٢	11	٠.٠٠٠٢	٢	1٥	٠.٠٠٠٤	٢	11	٠.٠٠٠٧	٢	11	٠.٠٠٠٧	٢	11	٠.٠٠٠٥	٢	11	٠.٠٠٠٥	٢	11	٠.٠٠٠٤
P٥٤	FLETCHCR	٥٠٠٠	(٠.٩٩, ..., ٠.٩٩)	٢	1٥	٠.٠٠1٧	٢	1٥	٠.٠٠٢٧	٢	11	٠.٠٠٢٢	٢	1٥	٠.٠٠٧٢	٢	1٥	٠.٠٠٨٢	٢	1٥	٠.٠٠٢٩	٢	11	٠.٠٠٥٥
P٥٥	EXT DENSCHNB	1٠٠	(٢, ..., ٢)	٩	٢٥	٠.٠٠٠٢	٨	٢٢	٠.٠٠1	٩	٢٩	٠.٠٠٥٥	٩	٢٥	٠.٠٠٠٧	٩	٢٥	٠.٠٠٠٧	٩	٢٥	٠.٠٠٠٤	٨	٢٢	٠.٠٠٠٤
P٥٦	EXT DENSCHNB	1٠٠٠	(٢, ..., ٢)	٩	٢٩	٠.٠٠٠٢	٨	٢٢	٠.٠٠٥٥	1٠	٢٢	٠.٠٠٥٥	٩	٢٥	٠.٠٠٠٩	٩	٢٥	٠.٠٠٠٩	٩	٢٩	٠.٠٠٠٥	٩	٢٢	٠.٠٠1٧
P٥٧	EXT DENSCHNB	1٠٠٠	(٢, ..., ٢)	٩	٢٩	٠.٠٠٠٥	٨	٢٢	٠.٠٠٠٧	1٠	٢٢	٠.٠٠٠٧	٩	٢٥	٠.٠٠11	٩	٢٥	٠.٠٠٠٧	٩	٢٩	٠.٠٠٠٩	٩	٢٨	٠.٠٠٠٨
P٥٨	EXT DENSCHNB	1٠٠٠٠	(٢, ..., ٢)	٩	٢٩	٠.٠٠٠٩	٨	٢٢	٠.٠٠1	1٠	٢٢	٠.٠٠٢٢	٩	٢٥	٠.٠٠٠٨	٩	٢٥	٠.٠٠1٢	٩	٢٩	٠.٠٠٠1	٩	٢٨	٠.٠٠٠1
P٥٩	EXT DENSCHNB	٢٠٠٠	(٢, ..., ٢)	٩	٢٩	٠.٠٠1٢	٨	٢٢	٠.٠٠٢	1٠	٢٢	٠.٠٠1٢	٩	٢٥	٠.٠٠1٧	٩	٢٥	٠.٠٠1٩	٩	٢٩	٠.٠٠1٧	٩	٢٨	٠.٠٠٢1
P٦٠	EXT DENSCHNB	٢٠٠٠	(٢, ..., ٢)	٩	٢٩	٠.٠٠1٧	٨	٢٢	٠.٠٠٥٥	1٠	٢٢	٠.٠٠٢٢	٩	٢٥	٠.٠٠٢٢	٩	٢٥	٠.٠٠٢٢	٩	٢٩	٠.٠٠1٨	٩	٢٨	٠.٠٠٢٥
P٦1	EXT DENSCHNB	1٠٠٠٠	(٢, ..., ٢)	٧	٢٥	٠.٠٠٢٥	٧	٢٥	٠.٠٠٢٢	1٠	٢٢	٠.٠٠٨1	٩	٢٥	٠.٠٠٥1	٩	٢٥	٠.٠٠٥٥	٧	٢٥	٠.٠٠٧٩	٩	٢٨	٠.٠٠٧
P٦٢	DIAGONAL ٢	1٠٠٠٠	(-1/١٠, ..., -1/١٠)	٢	11	٠.٠٠٠٩	٢	1٤	٠.٠٠٢٢	٢	1٠	٠.٠1٧٢	٢	1٠	٠.٠1٢٥	٢	1٠	٠.٠1٥٥	٢	11	٠.٠٠٢	٩	1٠	٠.٠1٢1

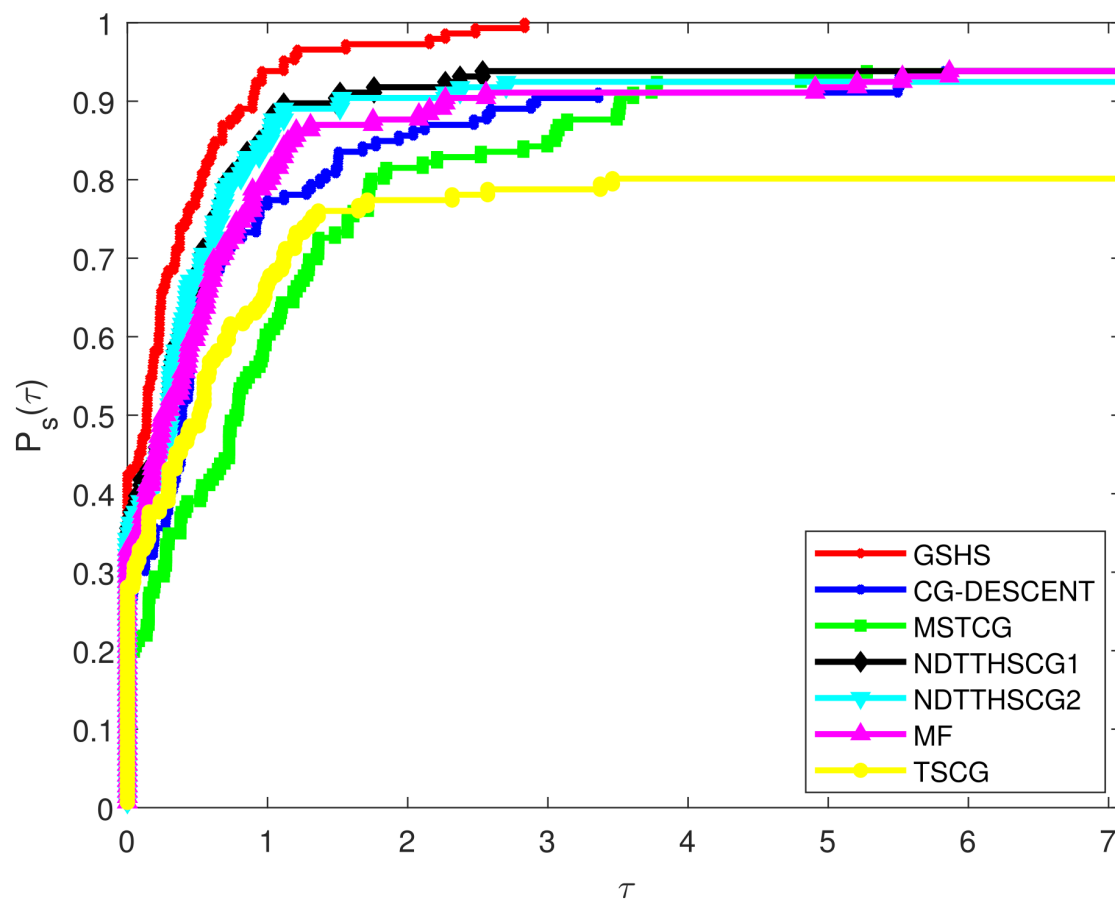
S/N	Test functions	Dimm	Initial point	GSHS			CG _D ESCENT			MSTCG			NDTHSCG\			NDTHSCG\			MF			TSCG		
				NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU
PV٢	DIAGONAL ٢	٥٠٠٠	-)١.٠١.....١.٠١(٢	١١	٠.٠٠٧	٣	١٢	٠.٠١٥٧	٢	١٠	٠.٠١٣٩	٢	١٠	٠.٠١٥٧	٢	١٠	٠.٠٠٩٩	٢	١١	٠.٠١١٨	٢	١٠	٠.٠٠٨٥	٢
PV٤	DIAGONAL ٢	١٠٠٠٠	-)١.٠٣.....١.٠٣(٢	١١	٠.٠١٥٥	٣	١٢	٠.٠١٥٢	٢	١٠	٠.٠١٣١	٢	١٠	٠.٠١٤٥	٢	١٠	٠.٠١٢٢	٢	١١	٠.٠١٧٢	٢	١٠	٠.٠١٣٢	٢
PV٥	DIAGONAL ٢	٥٠٠٠٠	-)١.٠٣.....١.٠٣(٢	١١	٠.٠٤١٨	٣	١٢	٠.٠٢٩٥	٢	١٠	٠.٠٥٥٨	٢	١٠	٠.٠٤٣٨	٢	١٠	٠.٠٢٩٢	٢	١١	٠.٠٦١٧	٢	١٠	٠.٠٤٣٢	٢
PV٦	DIAGONAL ٧	١٠	)١.....١(.	١٣	٠.٠٠٢٥	٣	١٣	٠.٠١٥٢	٣	١٣	٠.٠٠٢٩	٣	١٣	٠.٠٠٣٢	٣	١٣	٠.٠٠٣٢	٣	١٣	٠.٠٠٣٨	٣	١٣	٠.٠٠٤٥	٣
PV٧	DIAGONAL ٧	١٠٠	)١.....١(.	١٣	٠.٠٠٢٤	٣	١٣	٠.٠٠٥٨	٣	١٣	٠.٠٠٢٩	٣	١٣	٠.٠٠٢٣	٣	١٣	٠.٠٠٣٣	٣	١٣	٠.٠٠٥٧	٣	١٣	٠.٠٠٥٧	٣
PV٨	DIAGONAL ٧	١٠٠	)١.....١(.	١٣	٠.٠٠٢٣	٣	١٣	٠.٠٠٢٢	٣	١٣	٠.٠٠٢٨	٣	١٣	٠.٠٠٢٣	٣	١٣	٠.٠٠٢٣	٣	١٣	٠.٠٠٣٣	٣	١٣	٠.٠٠٤٧	٣
PV٩	DIAGONAL ٨	١٠٠	-)٠.٥.....٠.٥(	١٢	٠.٠٠٣	٣	١٢	٠.٠٠٢	٣	١٢	٠.٠٠٢٨	٣	١٢	٠.٠٠٢٩	٣	١٢	٠.٠٠٣٣	٣	١٢	٠.٠٠٢٣	٣	١٢	٠.٠٠٤٧	٣
PA٠	DIAGONAL ٨	٥٠٠	-)٠.٥.....٠.٥(	١٢	٠.٠٠٣٩	٣	١٢	٠.٠٠٢٣	٣	٢٢	٠.٠١١٣	٣	١٢	٠.٠٠٢٢	٣	١٢	٠.٠٠٢٣	٣	١٢	٠.٠٠٢٣	٣	١٢	٠.٠٠٤٧	٣
PA١	DIAGONAL ٩	١	)٣.....٣(	١٥١	٢٠	٠.٠٠٠٧	١٥٣	٢٥٧٥	٠.٣٠٢٢	***	***	***	٢٧٢	٢٠٥٢	٠.١٨١٥	***	***	***	٢٢٩	١٣٩٧	٠.١٥٢٨	***	***	***
PA٢	DENSCHINA	١٠٠٠	)١.....١(.	١٣	٢١	٠.٠٠٢٢	٩	٢٠	٠.٠٠٠٥	١٨	١٥٢	٠.٠٠١٢	٨	٢٥	٠.٠٠٠٧	٨	٢٥	٠.٠٠٠٨	١٢	٢٥	٠.٠٠٠٨	٨	٢١	٠.٠٠٠٧
PA٣	DENSCHINA	٣٠٠٠	)١.....١(.	١٢	٢٢	٠.٠٠٠٨	٩	٢٥	٠.٠٠٠٢	١٨	١٥٢	٠.٠٠١٢	٨	٢٥	٠.٠٠١٢	٨	٢٥	٠.٠٠٠٨	١٢	٢٥	٠.٠٠٠٨	٨	٢٩	٠.٠٠٠٩
PA٤	DENSCHINA	١٥٠٠٠	)١.....١(.	١٢	٥٢	٠.٠٠٢٢	٨	٢٥	٠.٠٠٢٢	١٨	١٥٢	٠.٠٠٢٢	٨	٢٥	٠.٠٠٢٢	٨	٢٥	٠.٠٠٢٢	١٢	٢٥	٠.٠٠٢٢	٩	٢٣	٠.٠٠٢٢
PA٥	DENSCHINA	٢٠٠٠٠	)١.....١(.	١٢	٥٢	٠.٠٠٢٨	٨	٢٥	٠.٠٠٣١	٢٠	١١٥	٠.٠٠٢٢	٨	٢٥	٠.٠٠٢٨	٨	٢٥	٠.٠٠٣١	١٢	٢٥	٠.٠٠٣١	٩	٢٣	٠.٠٠٣
PA٦	DENSCHINA	٥٠٠٠٠	)١.....١(.	١٢	٢٥	٠.٠٠٢٩	٨	٢٢	٠.٠٠٢٨	٢٠	١١٥	٠.٠٠٢٩	٨	٢٥	٠.٠٠٢٩	٨	٢٥	٠.٠٠٢٩	١٢	٢٥	٠.٠٠٢٩	٨	٢٣	٠.٠٠٢٢
PA٧	DENSCHINA	١٠٠٠٠٠	)١.....١(.	١٢	٥٢	٠.٠٠٢٩	٨	٢٢	٠.٠٠٢٩	٢٠	١١٥	٠.٠٠٢٩	٨	٢٥	٠.٠٠٢٩	٨	٢٥	٠.٠٠٢٩	١٢	٢٥	٠.٠٠٢٩	٩	٢٣	٠.٠٠٢٢
PA٨	DENSCHINC	٢	)١.....١(.	٢٠	٨٨	٠.٠٠٢٢	٢٢	١٩١	٠.٠٠٢٢	٢٥	١٠٠٠	٠.٠٠٢٢	٥٨	٢٢٧	٠.٠٠٢٢	٢٠	٢٥٢	٠.٠٠٢٢	٢٢	٢٢٧	٠.٠٠٢٢	٢٢	٢٢	٠.٠٠٢٢
PA٩	DENSCHINC	٢	)١.....١(.	٢٠	٨٨	٠.٠٠٢٢	٢٢	١٩١	٠.٠٠٢٢	٢٥	١٠٠٠	٠.٠٠٢٢	٥٨	٢٢٧	٠.٠٠٢٢	٢٠	٢٥٢	٠.٠٠٢٢	٢٢	٢٢٧	٠.٠٠٢٢	٢٢	٢٢	٠.٠٠٢٢
PA١٠	BLOCK DIAGONAL	١٠٠	)٠.٥.....٠.٥(	١٨	٧٥	٠.٠٠٢٢	٢٢	٩١	٠.٠٠٢٢	١٣	١٨٧	٠.٠٠٢٢	٢٠	٨٣	٠.٠٠٢٢	٢٠	٨٣	٠.٠٠٢٢	١٨	٧٥	٠.٠٠٢٢	٢٢	٢٢	٠.٠٠٢٢
PA١١	BLOCK DIAGONAL	١٠٠٠	)٠.٥.....٠.٥(	١٨	٧٥	٠.٠٠٢٢	٢٢	٩١	٠.٠٠٢٢	١٣	١٨٧	٠.٠٠٢٢	٢٠	٨٣	٠.٠٠٢٢	٢٠	٨٣	٠.٠٠٢٢	١٨	٧٥	٠.٠٠٢٢	٢٢	٢٢	٠.٠٠٢٢
PA١٢	BLOCK DIAGONAL	٥٠٠٠	)٠.٥.....٠.٥(	١٨	٧٥	٠.٠٠٢٢	٢٢	٩١	٠.٠٠٢٢	١٣	١٨٧	٠.٠٠٢٢	٢٠	٨٣	٠.٠٠٢٢	٢٠	٨٣	٠.٠٠٢٢	١٨	٧٥	٠.٠٠٢٢	٢٢	٢٢	٠.٠٠٢٢
PA١٣	BLOCK DIAGONAL	١٠٠٠٠	)٠.٥.....٠.٥(	١٨	٧٩	٠.٠٠١٣	٢٥	١٥٣	٠.٠٠٢	١٣	١٨٩	٠.٠٠١٨	٢٢	٩١	٠.٠٠١٧	٢٢	٩١	٠.٠٠١٧	١٨	٧٩	٠.٠٠١٧	٢٢	٢٢	٠.٠٠١٧
PA١٤	BLOCK DIAGONAL	٢٠٠٠٠	)٠.٥.....٠.٥(	١٨	٧٩	٠.٠٠١٦	٢٦	١٥٧	٠.٠٠٢٢	١٥	١٩٢	٠.٠٠٢	٢٢	٩١	٠.٠٠٢٣	٢٢	٩١	٠.٠٠٢٣	١٨	٧٩	٠.٠٠٢٣	٢٢	٢٢	٠.٠٠٢٣
PA١٥	BLOCK DIAGONAL	٥٠٠٠٠	)٠.٥.....٠.٥(	١٩	٧٩	٠.٠٠٢١	٢٦	١٥٧	٠.٠٠٢٢	١٥	٢٠٣	٠.٠٠٢٢	٢٢	٩١	٠.٠٠٢٣	٢٢	٩١	٠.٠٠٢٣	١٩	٧٩	٠.٠٠٢٣	٢٢	٢٢	٠.٠٠٢٣
PA١٦	BLOCK DIAGONAL	١٠٠٠٠٠	)٠.٥.....٠.٥(	٢٠	٨٣	٠.٠٠٢٧	٢٦	١٥٧	٠.٠٠٢٢	١٥	٢٠٥	٠.٠٠٢٧	٢٣	٩٥	٠.٠٠٢	٢٣	٩٥	٠.٠٠٢	٢٠	٨٣	٠.٠٠٢	٢٢	٢٢	٠.٠٠٢
PA١٧	DQDRIC	١٠٠٠	)٢.....٢(	٨٢	١٩٢٧	٠.٠٠٢	٢٦	١٥٥٥	٠.٠٠٢٢	٢٢٢	٧٢٢٩	٠.٠٠٢٢	***	***	***	***	***	***	٢٥٠	٨١٢٣	٠.٠٠٢	***	***	***
PA١٨	DQDRIC	٥٠٠٠	)٢.....٢(	٢٢	١٧٥٢	٠.٠٠٢	٢٦	٨١٢٣	٠.٠٠٢	٢٢٥	١١١٢٩	٠.٠٠٢	***	***	***	***	***	***	٥٧٩	١١٢٥٩	٠.٠٠٢	٢٢	٢٥	٠.٠٠٢
PA١٩	QUARTIC	١٠٠٠	)٢.....٢(	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢
P١٠٠	QUARTIC	٥٠٠٠	)٢.....٢(	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢
P١٠١	QUARTIC	١٠٠٠٠	)٢.....٢(	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢
P١٠٢	QUARTIC	٢٠٠٠٠	)٢.....٢(	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢
P١٠٣	QUARTIC	٥٠٠٠٠	)٢.....٢(	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢
P١٠٤	QUARTIC	١٠٠٠٠٠	)٢.....٢(	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢	٢	٣٥	٠.٠٠٢
P١٠٥	LINEAR PERTURBED	٥٠٠	)١٣.....١٣(	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢
P١٠٦	LINEAR PERTURBED	١٠٠٠٠	)١٣.....١٣(	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢
P١٠٧	LINEAR PERTURBED	٥٠٠٠٠	)١٣.....١٣(	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢
P١٠٨	LINEAR PERTURBED	١٠٠٠٠٠	)١٣.....١٣(	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢	٢	٩	٠.٠٠٢
P١٠٩	ENGVAL١	٢٠	-)١.....١(	٣٥	١٢٦	٠.٠٠٢	٢٢	١٥٣	٠.٠٠٢	٢٦	١٢٩	٠.٠٠٢	٢٦	١١٥	٠.٠٠٢	٢٦	١١٥	٠.٠٠٢	٢٥	١٠٥	٠.٠٠٢	***	***	***

جدول ٣.٢: نتایج عددی الگوریتمها

S/N	Test functions	Dimm	Initial point	GSHS			CG _D ESCENT			MSTCG			NDTHSCG\			NDTHSCG\			MF			TSCG		
				NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU	NI	FE	CPU
P110	TRIG WHITE-HOLST	50	-)1.....1(	29	121	0.0002	22	153	0.0002	20	200	0.0002	26	115	0.0002	26	115	0.0002	27	113	0.0002	***	***	***
P111	ENGVAL1	100	-)1.....1(	29	121	0.0002	22	153	0.0002	20	219	0.0002	26	114	0.0002	26	114	0.0002	27	113	0.0002	***	***	***
P112	ENGVAL1	200	-)1.....1(	28	117	0.0002	25	157	0.0002	22	189	0.0002	27	114	0.0002	***	***	***	35	125	0.0002	***	***	***
P113	ENGVAL1	300	-)1.....1(	21	129	0.0002	21	129	0.0002	***	***	***	27	114	0.0002	27	114	0.0002	21	129	0.0002	***	***	***
P114	ENGVAL1	400	-)1.....1(	23	147	0.0002	23	147	0.0002	***	***	***	27	115	0.0002	***	***	***	***	***	***	***	***	***
P115	ENGVALA	10	2.....2(	22	178	0.0002	25	278	0.001	22	235	0.0002	21	247	0.0002	21	247	0.0002	22	221	0.0002	28	223	0.0002
P116	ENGVALA	100	2.....2(	112	223	0.0002	26	278	0.0002	41	278	0.0002	***	***	***	22	202	0.0002	25	272	0.0002	28	273	0.0002
P117	DENSCHNF	10000	-)10.....10(	13	182	0.0002	14	10	0.0002	12	9	0.0002	12	26	0.001	12	26	0.0002	13	108	0.0002	18	102	0.0002
P118	DENSCHNF	100000	-)10.....10(	13	182	0.001	12	9	0.001	12	9	0.001	12	26	0.001	12	26	0.001	13	112	0.001	12	178	0.001
P119	DENSCHNF	1000000	-)10.....10(	13	182	0.002	12	10	0.002	12	9	0.002	12	26	0.002	12	26	0.002	13	112	0.002	12	178	0.002
P120	DENSCHNF	10000000	-)10.....10(	14	182	0.01	15	10	0.01	12	9	0.01	12	26	0.01	12	26	0.01	13	112	0.01	11	178	0.01
P121	DENSCHNF	100000000	-)10.....10(	14	182	0.05	15	10	0.05	12	9	0.05	12	26	0.05	12	26	0.05	13	112	0.05	11	178	0.05
P122	DENSCHNF	200000000	-)10.....10(	13	182	0.1	15	10	0.1	12	9	0.1	12	26	0.1	12	26	0.1	13	112	0.1	9	178	0.1
P123	ROTATED ELLIPSE Y	Y	1)1.(	1	26	0.0002	1	20	0.0002	1	20	0.0002	1	20	0.001	1	20	0.0002	1	20	0.0002	1	20	0.0002
P124	SIX HUMP	Y	2)2,2,2.(	10	26	0.001	10	20	0.0002	12	25	0.0002	11	27	0.0002	11	27	0.0002	10	26	0.0002	12	102	0.0002
P125	SIX HUMP	Y	2)2,2,2.(	11	102	0.0002	11	20	0.0002	12	102	0.0002	15	102	0.0002	15	102	0.0002	11	102	0.0002	18	171	0.0002
P126	SIX HUMP	Y	2)2,2,2.(	11	189	0.0002	12	10	0.0002	12	102	0.0002	15	102	0.0002	15	102	0.0002	11	102	0.0002	20	181	0.0002
P127	PRICE f	Y	-)0.2,-0.2(	2	10	0.0002	2	10	0.0002	2	11	0.0002	2	10	0.0002	2	10	0.0002	2	10	0.0002	2	11	0.0002
P128	PRICE f	Y	-)0.1,-0.1(	2	9	0.0002	2	9	0.0002	2	10	0.0002	2	10	0.0002	2	9	0.0002	2	9	0.0002	2	10	0.0002
P129	ALUFFI PENTINI	Y	1)1.(	4	12	0.0002	4	12	0.0002	4	18	0.0002	4	12	0.0002	4	12	0.0002	4	12	0.0002	4	12	0.0002
P130	ALUFFI PENTINI	Y	2)2.(	3	12	0.0002	3	12	0.0002	3	20	0.0002	3	21	0.001	3	21	0.0002	3	12	0.0002	4	21	0.0002
P131	EXT HIMMELAU	100000	1)1.....1(	26	204	0.22	27	218	0.22	27	222	0.22	28	222	0.22	28	222	0.22	27	222	0.22	28	222	0.22
P132	EXT HIMMELAU	200000	1)1.....1(	28	202	0.42	28	218	0.42	28	222	0.42	28	222	0.42	28	222	0.42	28	222	0.42	28	222	0.42
P133	EXT HIMMELAU	400000	1)1.....1(	25	218	0.32	27	218	0.32	27	222	0.32	27	222	0.32	27	222	0.32	27	222	0.32	27	222	0.32
P134	EXT HIMMELAU	1000000	1)1.....1(	27	212	0.22	28	212	0.22	28	222	0.22	28	222	0.22	28	222	0.22	28	222	0.22	28	222	0.22
P135	ROTATED ELLIPSE Y	Y	1)1.....0.01(	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002
P136	ROTATED ELLIPSE Y	Y	2)0.2,-0.2(	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002
P137	ROTATED ELLIPSE Y	Y	2)2,2.(	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002
P138	ROTATED ELLIPSE Y	Y	2)2,2,2.(	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002	22	282	0.0002
P139	EXT TRIDIAGONAL 1	1000	2).....2(	A	21	0.0002	A	22	0.0002	20	22	0.0002	19	22	0.0002	19	22	0.0002	A	21	0.0002	21	22	0.0002
P140	EXT TRIDIAGONAL 1	10000	2).....2(	A	21	0.0002	A	22	0.0002	20	22	0.0002	19	22	0.0002	19	22	0.0002	A	21	0.0002	21	22	0.0002
P141	EXT TRIDIAGONAL 1	100000	2).....2(	A	21	0.0002	A	22	0.0002	20	22	0.0002	19	22	0.0002	19	22	0.0002	A	21	0.0002	21	22	0.0002
P142	EXT TRIDIAGONAL 1	200000	2).....2(	2	22	0.0002	10	20	0.0002	20	22	0.0002	19	22	0.0002	19	22	0.0002	2	22	0.0002	10	22	0.0002
P143	EXT TRIDIAGONAL 1	400000	2).....2(	2	22	0.0002	10	20	0.0002	20	22	0.0002	19	22	0.0002	19	22	0.0002	2	22	0.0002	10	22	0.0002
P144	EXT TRIDIAGONAL 1	1000000	2).....2(	2	22	0.0002	10	20	0.0002	20	22	0.0002	19	22	0.0002	19	22	0.0002	2	22	0.0002	10	22	0.0002
P145	THREE HUMP	Y	2)0.2,0.2(	15	28	0.0002	15	28	0.0002	15	111	0.0002	17	28	0.0002	17	28	0.0002	15	28	0.0002	22	28	0.0002
P146	THREE HUMP	Y	1)0.2,(	15	28	0.0002	***	***	***	11	112	0.0002	22	28	0.0002	22	28	0.0002	15	28	0.0002	22	28	0.0002



شکل ۱.۲



شکل ۲.۲

بنابراین، بهترین روش توسط نمودار پروفایل عملکرد در منحنی بالا نمایش داده شده است. از این رو، نتایج آزمایش‌ها به صورت گرافیکی در شکل‌های ۱-۳ ارائه شده‌اند که نشان می‌دهند روش GSHS کارآمد است و بر اساس سه معیار، نسبت به سایر روش‌ها عملکرد بهتری دارد.

در شکل ۱، روش GSHS به نرخ برد ۵۵٪ دست یافته است و ۱۰۰٪ مسائل را با موفقیت حل کرده است. روش‌های TSCG, MSTCG, NDTTHSCG1 و NDTTHSCG2 به ترتیب با نرخ برد حدود ۴۲٪، ۴۱٪، ۳۹٪ و ۳۸٪ در رتبه‌های بعدی قرار دارند. با این حال، روش TSCG علی‌رغم داشتن نرخ برد بالاتر نسبت به روش‌های MF و کاهش CG - کمترین نرخ حل مسئله را دارد که ۷۸٪ است. در حالی که هر دو روش MF و کاهش CG - نرخ ۹۳.۸٪ را ثبت کرده‌اند. شکل ۲ نشان می‌دهد که GSHS عملکرد بهتری نسبت به روش‌های NDTTHSCG1, NDTTHSCG2 و MF داشته است و با نرخ حدود ۴۲٪ برد تمامی مسائل را با موفقیت حل کرده است. در مقابل، روش‌های TSCG، کاهش CG و MSTCG کمترین میزان موفقیت را داشته‌اند، با نرخ برد به ترتیب ۲۹٪، ۲۸٪ و ۲۰٪. شکل ۳ عملکرد روش‌ها را بر اساس زمان پردازش نمایش می‌دهد. مشخص است که روش پیشنهاد شده (GSHS) با ۵۹٪ برد برتری قابل توجهی نسبت به سایر روش‌ها داشته است در حالی که روش‌های کاهش CG، TSCG, MF, NDTTHSCG2، NDTTHSCG1، MSTCG تنها حدود ۱۰٪ برد داشته‌اند. بنابراین، شکل‌های ۱-۳ نشان می‌دهند که روش GSHS در مقایسه با سایر روش‌های گرادیان مزدوج برتر و کارآمدتر است.

۵.۲ کاربردهای GSHS در حل مسائل نوسازی تصویر

روش‌های گرادیان مزدوج غیرخطی به عنوان الگوریتم‌های عددی کارآمد و به صورت سراسری همگرا شناخته می‌شوند. این روش‌ها نه تنها برای مسائل بهینه‌سازی غیرساختاری بلکه روی مسائل کاربردی دنیای واقعی موثر هستند. در سال‌های اخیر، مطالعات زیادی مناسب بودن روش‌های گرادیان مزدوج مختلف را بر روی مسائل کاربردی مختلف مورد بررسی قرار داده‌اند از جمله، کنترل حرکت رباتیک [۴۲]، انتخاب سبد سرمایه‌گذاری [۳۱]، تحلیل رگرسیون [۲۳، ۴۷]، شبکه عصبی [۲۵] و رفع تاری تصویر [۴۸، ۴۶]. مسائل رفع تاری تصویر به دلیل اهمیت در حوزه‌های سلامت و امنیت، به طور گسترده در پژوهش‌ها مورد بررسی قرار گرفته‌اند. این مسائل شامل بازسازی تصاویر نویزی یا تار شده هستند و انتخاب یک روش مناسب برای این کار همواره چالشی اساسی بوده است. در این گزارش، هدف بررسی عملکرد روش پیشنهادی مبتنی بر گرادیان در مسائل بازسازی تصویر از فرم زیر است که ϕ_α در $\mathcal{H}(u)$ به عنوان تابع پتانسیل حفظ‌کننده لبه تعریف می‌شود $\phi_\alpha(t) = \sqrt{t^2 + \alpha}$

که در آن $\alpha = 1$ در نظر گرفته شده است [۱۷]

$$\mathcal{H}(u) = \sum_{(i,j) \in G} \left\{ \sum_{(m,n) \in T_{i,j} \setminus G} \phi_{\alpha}(u_{i,j} - \xi_{m,n}) + \frac{1}{\gamma} \sum_{(m,n) \in T_{i,j} \cap G} \phi_{\alpha}(u_{i,j} - u_{m,n}) \right\},$$

برای اعمال روش پیشنهادی، باید مسئله بازسازی تصویر را به مسئله بهینه‌سازی $\min \mathcal{H}(u)$ تبدیل شود که در آن، تصویر اصلی x شامل $M \times N$ پیکسل است و G مجموعه‌ای از نویزهای کاندید x را تعریف می‌کند که مقادیر آن به صورت زیر تعیین می‌شود

$$G = \{(i, j) \in \mathbb{Q} \mid \bar{\xi}_{ij} \neq \xi_{ij}, \xi_{ij} = s_{\min} \text{ or } s_{\max}\}, \quad (53.2)$$

اصطلاحات موجود در معادله (53.2) عبارتند از: $s_{\max}$ نشان‌دهنده بیشترین مقدار نویز پیکسلی است در حالی که $s_{\min}$ کمترین مقدار نویز پیکسلی را نشان می‌دهد. برای هر $i, j \in \mathbb{Q} = \{1, 2, \dots, M\} \times \{1, 2, \dots, N\}$ همسایگی ذکر شده به صورت $T_{ij} = \{(i, j-1), (i, j+1), (i-1, j), (i+1, j)\}$ تعریف می‌شود. همچنین، $\bar{\xi}$ فیلتر میانگین تطبیقی از تصویر نویزدار مشاهده شده ξ است.

برای هدف این مطالعه، دو تصویر (PEA (512 × 512) و (PAPAYA (512 × 512) BUTTERFLY) که توسط نویز ضربه‌ای نمک و فلفل تخریب شده‌اند برای بازسازی در نظر گرفته می‌شوند. این مطالعه، کیفیت تصویر بازسازی شده را بر اساس سه معیار ارزیابی می‌کند، نسبت اوج سیگنال به نویز، (PSNR)، خطای نسبی (RelErr) و زمان پردازش (CPU). PSNR یکی از پرکاربردترین معیارهای ارزیابی کیفیت بصری تصویر، در بازسازی تصویر است. این معیار، حداکثر توان ممکن سیگنال را نسبت به توان نویزهای مخربی که بر دقت بازسازی تاثیر می‌گذارند محاسبه می‌کند. PSNR معیار کلیدی برای ارزیابی کیفیت تصاویر بازسازی شده در مقایسه با تصویر اصلی است. هر چه مقدار PSNR بالاتر باشد، کیفیت تصویر خروجی بهتر خواهد بود [۳۴]. بنابراین، روش‌هایی که مقادیر PSNR بالاتری دارند، بسیار کارآمد در نظر گرفته می‌شوند. PSNR به صورت زیر محاسبه می‌شود. در اینجا MAX_I مقدار بیشینه پیکسل ممکن در تصاویر را تعریف می‌کند و خطای مربعی میانگین MSE، تفاوت بین مقادیر میانگین پیکسل‌های تصاویر کامل را اندازه‌گیری می‌کند

$$\begin{aligned} PSNR &= 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \\ &= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE). \end{aligned}$$

هر چه مقدار MSE بیشتر باشد، تفاوت بین تصاویر بازسازی شده و تخریب شده بیشتر است. با این حال،

روش		خروجی‌های بازسازی تصویر بر اساس CPU			
تصویر		GSHS	NDTTHSCG1	$CG_{DESCENT}$	MSTCG
PAPAYA	۴۰٪	۳۲.۶۲۲۸	۳۲.۵۵۵۶	۳۴.۰۴۵۵	۳۲.۵۶۷۲
PAPAYA	۸۰٪	۶۷.۰۳۲۹	۸۶.۷۲۸۲	۸۸.۵۷۳۱	۸۶.۶۰۸۳
BUTTERFLY PEA	۴۰٪	۴۷.۰۲۱۵	۳۱.۶۶۷۱	۳۲.۶۸۷۸	۳۲.۶۰۳۴
BUTTERFLY PEA	۸۰٪	۹۴.۷۵۵۳	۹۲.۱۲۲۱	۸۸.۲۷۳۴	۱۱۰.۹۶۸۲

Table ۲.۵: پردازش زمان اساس بر تصویر بازسازی خروجی‌های: جدول ۵

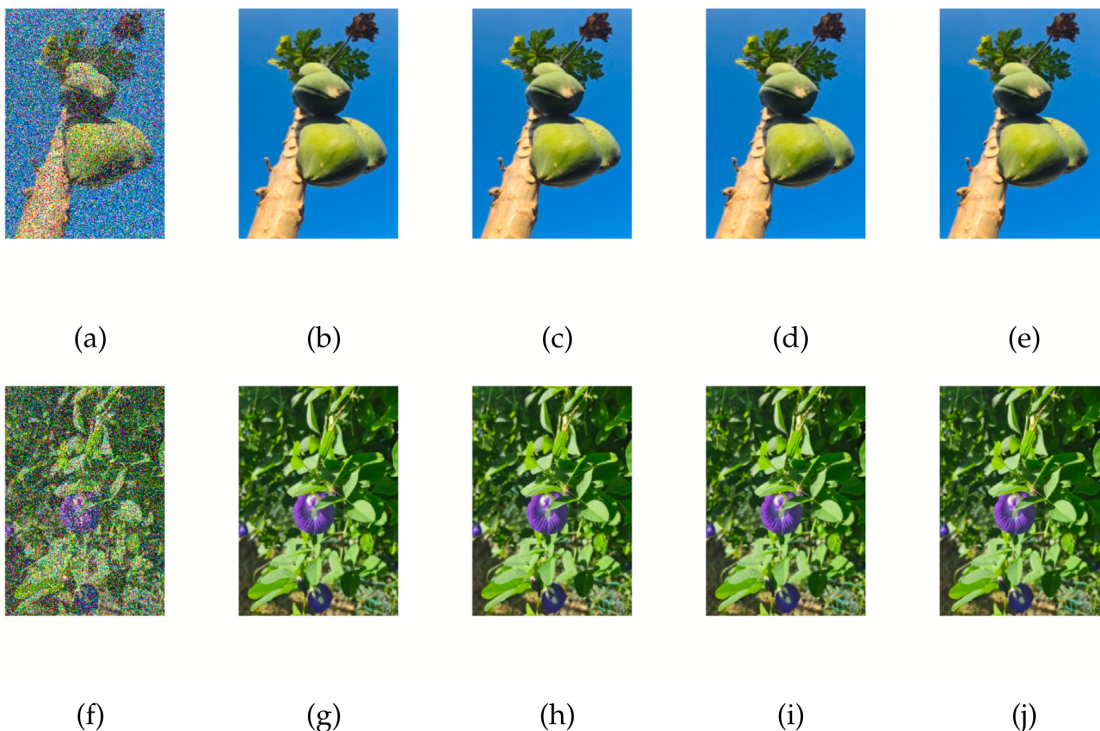
روش		خروجی‌های بازسازی تصویر بر اساس خطای نسبی	
تصویر		GSHS	NDTTHSCG1
PAPAYA	۴۰٪	۱.۰۱۹۶	۱.۰۳۷۹
PAPAYA	۸۰٪	۱.۴۴۲۶	۱.۴۲۳۷
BUTTERFLY PEA	۴۰٪	۲.۳۰۷۱	۱.۴۶۱۲
BUTTERFLY PEA	۸۰٪	۳.۹۷۱۷	۳.۸۸۷۲

Table ۲.۶: (RelErr) نسبی خطای اساس بر تصویر بازسازی خروجی‌های: جدول ۶

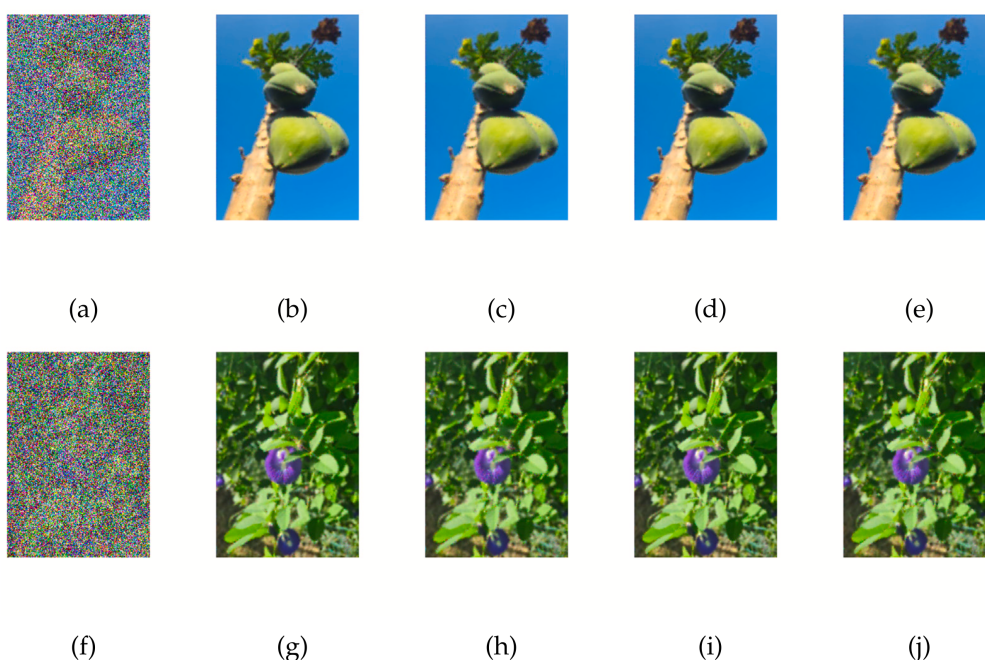
روش		PSNR خروجی‌های بازسازی تصویر بر اساس			
تصویر		GSHS	NDTTHSCG1	$CG_{DESCENT}$	MSTCG
PAPAYA	۴۰٪	۲۶.۸۱۴۴	۲۶.۰۱۷۸	۲۶.۰۲۵۶	۲۵.۹۳۹۹
PAPAYA	۸۰٪	۲۳.۳۳۳۶	۱۸.۷۲۵۸	۱۹.۵۲۶۵	۱۸.۳۵۲۲
BUTTERFLY PEA	۴۰٪	۲۰.۷۹۳۰	۲۰.۵۷۳۴	۲۰.۶۷۵۱	۲۰.۶۳۵۳
BUTTERFLY PEA	۸۰٪	۱۵.۵۷۶۳	۱۳.۰۸۶۶	۱۴.۷۷۴۲	۱۴.۰۲۷۰

جدول ۷.۲: جدول ۷: خروجی‌های بازسازی تصویر بر اساس PSNR

هنگام بررسی لبه‌ها باید دقت بیشتری به خرج داد. MSE به صورت $MSE = \frac{1}{N} \sum \sum (E_{i,j} - o_{i,j})^2$ محاسبه می‌شود که در آن N اندازه تصویر اصلی o است و E نمایانگر تصویر لبه می‌باشد. نتایج زیر عملکرد روش پیشنهادی را در مقایسه با سایر روش‌ها با ویژگی‌های مشابه نشان می‌دهند. جداول ۵، ۶ و ۷ عملکرد تمامی الگوریتم‌ها را بر اساس زمان پردازش، خطای نسبی (RelErr) و PSNR نمایش می‌دهند. این معیارها برای اندازه‌گیری کارایی و پایداری الگوریتم پیشنهادی انتخاب شده‌اند. بررسی دقیق نتایج نشان می‌دهد اگرچه روش پیشنهادی در تمامی معیارها از سایر الگوریتم‌های موجود پیشی نگرفت، اما عملکرد برتری از خود نشان داده‌است. این برتری در کیفیت بهتر تصاویر خروجی با کمترین زمان پردازش و مقادیر بالاتر PSNR مشهود است. بر اساس این نتایج، نتیجه گرفته می‌شود که الگوریتم جدید ما بهترین عملکرد را دارد و تصاویر با کیفیت بالاتری را بازسازی می‌کند همان طور که در شکل‌های ۴ و ۵ نشان داده شده‌است. این مقایسه با سایر الگوریتم‌های مشابه همچنین پایداری الگوریتم جدید را در حل مسائل کاربردی دنیای واقعی نشان می‌دهد.



شکل ۳.۲: تصاویر PAPAYA و BUTTERFLY PEA که با ۴۰٪ نویز نمکی-فلغلی خراب شده‌اند: (a,f)، تصاویر بازیابی‌شده با استفاده از GSHS: (b,g)، NDTTHSCG1: (c,h)، CG-DESCENT: (d,i)، MSTCG: (e,j)



شکل ۴.۲: تصاویر PAPAYA و BUTTERFLY PEA که با ۸۰٪ نویز نمکی-فلغلی خراب شده‌اند: (a,f)، تصاویر بازیابی‌شده با استفاده از GSHS: (b,g)، NDTTHSCG1: (c,h)، CG-DESCENT: (d,i)، MSTCG: (e,j)

نتیجه‌گیری

در این گزارش، با در نظر گرفتن شرط هم‌ارزی تعمیم‌یافته و همچنین ویژگی همگرایی مناسب روش نیوتن یک روش گرادیان مزدوج طیفی برای مسائل بهینه‌سازی غیرساختاری با ابعاد بزرگ بر اساس شرط دایه و لیا و ارائه شده‌است. پارامتر دایه و لیا و t از منبع [۹] برای دستیابی به همگرایی سراسری اتخاذ شده‌است و آزمایش‌های عددی برای روش پیشنهادی اجرا شده‌است. با اعمال تکنیک راه‌اندازی مجدد پاول [۳۷] ویژگی نزولی کافی و همگرایی بدون اعمال خاصیت دوگانه برش یافته بر روی روش HS حاصل شده‌است. برخی گزارشات عددی بر روی مسائل غیرخطی استاندارد و همچنین مسائل بازسازی تصویر نشان می‌دهند که روش پیشنهادی در مقایسه با TSCG، NDTTHSCG1، NDTTHSCG2، گرادیان مزدوج کاهشی، MSTCG و روش‌های MF کارا تر است. به عنوان کاری در آینده، باور بر این است که امکان حذف کامل خاصیت کران‌دار اعمال شده در [۱، ۱۴، ۲۴، ۴۸] بر اصلاحات گرادیان مزدوج وجود دارد و همگرایی و ویژگی نزولی کافی آنها تضمین شود.

کتاب نامه

- [1] K. Amini and P. Faramarzi, (2023), Global convergence of a modified spectral three-term CG algorithm for nonconvex unconstrained optimization problems, *J. Comput. Appl. Math.* 417, Article ID 114630, 12 p.
- [2] Z. Aminifard and S. Babaie-Kafaki, (2019), A modified descent Polak–Ribière–Polyak conjugate gradient method with global convergence property for nonconvex, *functions Calcolo* 56, No. 2, Paper No. 16, 11 p.
- [3] N. Andrei, (2020), Nonlinear conjugate gradient methods for unconstrained optimization, *Cham: Springer*
- [4] S. Babaie-Kafaki, (2023), A survey on the Dai–Liao family of nonlinear conjugate gradient methods *RAIRO, Oper. Res.* 57 No. 1, 43–58.
- [5] S. Babaie-Kafaki and R. Ghanbari (2015), Two optimal Dai-Liao conjugate gradient methods *Optimization* 64, 2277–2287.
- [6] J. Barzilai and J. M. Borwein (1988), Two-point step size gradient methods, *IMA J. Numer. Anal.* 8, No. 1, 141–148.
- [7] E. G. Birgin and J. M. Martínez, (2001), A spectral conjugate gradient method for unconstrained optimization, *Appl. Math. Optim.* 43, No. 2, 117–128.
- [8] F. Chen et al, (2023), A descent modified hs conjugate gradient method with an optimal property, *RAIRO, Oper. Res.* 57, No. 2, 541–550.
- [9] Y.-H. Dai and C.-X. Kou, (2013), A nonlinear conjugate gradient algorithm with

- an optimal property and an improved Wolfe line search, *SIAM J. Optim.* 23, No. 1, 296–320.
- [10] Y. H. Dai and L. Z. Liao, (2001), New conjugacy conditions and related nonlinear conjugate gradient methods, *Appl. Math. Optim.* 43, No. 1, 87–101.
- [11] Y. Dai and Y. Yuan, (1999), A nonlinear conjugate gradient method with a strong global convergence property, *SIAM J. Optim.* 10, No. 1, 177–182.
- [12] F. E. D. Dolan and J. J. Moré, (2002), Benchmarking optimization software with performance profiles *Math. Program.* 91, No. 2 (A), 201–213.
- [13] P. Faramarzi and K. Amini, (2009), A modified spectral conjugate gradient method with global convergence, *J. Optim. Theory Appl.* 182, No. 2, 667–690.
- [14] P. Faramarzi and K. Amini, (2021), A spectral three-term Hestenes–Stiefel conjugate gradient method, *4OR* 19, No. 1, 71–92.
- [15] R. Fletcher, (1987), Practical methods of optimization, 2nd ed. Chichester: Wiley.
- [16] R. Fletcher and C. M. Reeves, (1964), Function minimization by conjugate gradients, *Comput. J.* 7, 149–154.
- [17] G. Yu et al, (2010), A descent spectral conjugate gradient method for impulse noise removal, *Appl. Math. Lett.* 23, No. 5, 555–560.
- [18] W. W. Hager and H. Zhang, (2005), A new conjugate gradient method with guaranteed descent and an efficient line search, *SIAM J. Optim.* 16, No. 1, 170–192.
- [19] W. W. Hager and H. Zhang, (2006), Cg descent, a conjugate gradient method with guaranteed descent, *ACM Trans. Math. Softw.* 32, No. 1, 113–137.
- [20] W. W. Hager and H. Zhang, (2006), A survey of nonlinear conjugate gradient methods, *Pac. J. Optim.* 2, No. 1, 35–58.

- [21] A. S. Halilu et al, (2021), Signal recovery with convex constrained nonlinear monotone equations through conjugate gradient hybrid approach, *Math. Comput. Simul.* 187, 520–539.
- [22] M. R. Hestenes and E. Stiefel, (1952), Methods of conjugate gradients for solving, *J. Res. Natl. Bur. Stand.* 49, 409–436.
- [23] I. M. Sulaiman and M. Mamat, (2020), A new conjugate gradient method with descent properties and its application to regression analysis, *JNAIAM, J. Numer. Anal. Ind. Appl. Math.* 14, No, 1–2, 25–39.
- [24] J. Jian et al, (2017), A new spectral conjugate gradient method for large-scale unconstrained optimization, *Optim. Methods Softw.* 32, No. 3, 503–515.
- [25] K. Kamilu, M. Sulaiman, A. Muhammad, A. Mohamad, M. Mamat (2023), *Mathematical Modeling and Computing*, Vol. 10, No. 2, pp, 326–337.
- [26] M. Khoshsimaye-Bargard and A. Ashrafi, (2023), A family of the modified three-term Hestenes–Stiefel conjugate gradient method with sufficient descent and conjugacy conditions, *J. Appl. Math. Comput.* 69, No. 3, 2331–2360.
- [27] D. Li and M. Fukushima, (2001) A modified BFGS method and its global convergence in nonconvex minimization, *J. Comput. Appl. Math.* 129, No, 1–2, 15–35.
- [28] J. Liu and Y. Feng, (2019), A derivative-free iterative method for nonlinear monotone equations with convex constraints, *Numer. Algorithms* 82, No. 1, 245–262.
- [29] J. K. Liu and S. J. Li, (2014), New hybrid conjugate gradient method for unconstrained optimization, *Appl. Math. Comput.* 245, 36–43.
- [30] Y. Liu and C. Storey, (1991), Efficient generalized conjugate gradient algorithms, *J. Optim. Theory Appl.* 69, No. 1, 129–137.
- [31] M. Maulana, M.S. Ibrahim, B.A. Auwal, A. Gianinna, F. Sukon (2023), A new family of hybrid three-term conjugate gradient method for unconstrained optimiza-

- tion with application to image restoration and portfolio selection, *AIMS Mathematics*, 1–28.
- [32] J. Momin, Y. Xin-Sh (2013), A literature survey of benchmark functions for global optimization problems, *Int. J. Math. Model. Numer. Optim.* 4 (2), 150–194.
- [33] H. Mrad and S. M. Fakhari, (2024), Optimization of unconstrained problems using a developed algorithm of spectral conjugate gradient method calculation, *Math. Comput. Simul.* 215, 282–290.
- [34] M. Nadipally, (2019), Optimization of methods for image-texture segmentation using ant colony optimization, *Intelligent Data Analysis for Biomedical Applications*, Elsevier, 21–47.
- [35] E. Polak and G. Ribière, (1969), Note sur la convergence de méthodes de directions conjuguées, *Rev. Franç. Inform. Rech. Opér.* 3, No. 16, 35–43.
- [36] B.T. Polyak, (1967), A general method for solving extremal problems, *Dokl. Akad. Nauk SSSR* 174 (1), 33–36.
- [37] M. Powell, (1984), Nonconvex minimization calculations and the conjugate gradient method, in: D.F. Griffiths (Ed.) Numerical Analysis, in: Lecture Notes in, *Mathematics*, vol. 1066, 122–141.
- [38] M. Raydan, (1997) The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem, *SIAM J. Optim.* 7 (1), 26–33.
- [39] M. Rivaie, M. Mamat, L.W. June, I. Mohd, (2012), A new class of nonlinear conjugate gradient coefficients with global convergence properties, *Appl. Math*, 11323–11332.
- [40] N. Salihu, H.A. Babando, I. Arzuka, S. Salihu, (2023), A hybrid conjugate gradient method for unconstrained optimization with application, *Bangmod Int. J. Math. Comput. Sci.* 9, 24–44.

- [41] N. Salihu, P. Kumam, A.M. Awwal, I. Arzuka, T. Seangwattana, (2023), A structured Fletcher-Reeves spectral conjugate gradient method for unconstrained optimization with application in robotic model, *Oper. Res. Forum* .
- [42] N. Salihu, P. Kumam, A.M. Awwal, I.M. Sulaiman, T. Seangwattana, (2023), The global convergence of spectral RMIL conjugate gradient method for unconstrained optimization with applications to robotic model and image recovery, *PLoS One* 18 (3).
- [43] N. Salihu, P. Kumam, I.M. Sulaiman, T. Seangwattana, (2023), An efficient spectral minimization of the Dai-Yuan method with application to image reconstruction, *AIMS Math.* 8 (12), 30940—30962.
- [44] N. Salihu, P. Kumam, M.Muhammad. Yahaya, T. Seangwattana, (2024), A revised Liu–Storey conjugate gradient parameter for unconstrained optimization problems, *with applications, Eng. Optim*, 1—25.
- [45] N. Salihu, M.R. Odekunle, A.M. Saleh, S. Salihu, (2021), A Dai-Liao hybrid Hestenes-Stiefel and Fletcher-Reeves methods for unconstrained optimization, *Int. J. Ind. Optim.* 2 (1), 33—50.
- [46] B.-K. Saman, M. Nasrin, A. Zohre, (2023), A descent extension of a modified Polak-Ribière-Polak method with application in image restoration problem, *Optim, Lett.* 17 (6), 1—17.
- [47] I.M. Sulaiman, M. Malik, A.M. Awwal, P. Kumam, M. Mamat, S. Al-Ahmad, (2022), On three-term conjugate gradient method for optimization problems with applications on COVID-19 model and robotic motion control, *Adv. Contin. Discrete Models*, 1—22.
- [48] X. Wang, (2023), A class of spectral three-term descent Hestenes-Stiefel conjugate gradient algorithms for large-scale unconstrained optimization and image restoration problems, *Appl. Numer. Math.* 192, 41—56.

- [49] M.M. Yahaya, P. Kumam, A.M. Awwal, P. Chaipunya, S. Aji, S. Salisu, (2022), A new generalized Quasi-Newton algorithm based on structured diagonal Hessian approximation for solving nonlinear least-squares problems with application to 3DoF planar robot arm manipulator, *IEEE Access* 10, 10816–10826.
- [50] L. Zhang, W. Zhou, D.-H. Li, (2006), A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence, *IMA J. Numer. Anal.* 26 (4), 629–640.
- [51] L. Zhang, W. Zhou, D. Li, (2006), Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search, *Numer. Math.* 104(4), 561–572.
- [52] L. Zhang, W. Zhou, D. Li, (2007), Some descent three-term conjugate gradient methods and their global convergence, *Optim. Methods Softw.* 22 (4), 697–711.
- [53] G. Zoutendijk, (1970), Nonlinear programming, computational methods, in: J. Abadie (Ed.), Integer and Nonlinear Programming, *North-Holland, Amsterdam*, 37–86.

واژه‌نامه

ا

step size اندازه گام
portfolio selection انتخاب سبد سرمایه‌گذاری

ب

Unconstrained optimization بهینه‌سازی غیرساختاری
Image reconstruction بازسازی تصویر
memoryless بدون حافظه
maximum noisy pixel بیشترین مقدار نویز پیکسلی

پ

spectral parameter پارامتر طیفی
scalar parameter پارامتر اسکالر
jamming phenamana پدیده جمینگ (گیرکردگی)
bounded below پایین کران‌دار
performance profile پروفایل عملکرد
robustness پایداری

ت

objective function تابع هدف
Powell's restart strategy تکنیک راه‌اندازی مجدد پاول
quadratic function تابع درجه دوم

line search technique..... تکنیک جست‌وجوی خطی
 regression analysis تحلیل رگرسیون

ج

search direction جهت جستجو
 sufficient descent direction..... جهت نزولی کافی
 exact line search..... جست‌وجوی خطی دقیق
 inexact line search جست‌وجوی خطی تقریبی
 Newton direction..... جهت نیوتن
 strong Wolfe line search..... جست‌وجوی خطی ولف قوی
 optimal search direction جهت جست‌وجوی بهینه

ح

solver..... حل‌کننده

خ

double bounded property خاصیت کران‌دار بودن
 quadratic convergence property..... خاصیت همگرایی درجه دوم
 descent property خاصیت نزولی
 self-scaling خودمقیاس‌دهی
 local property خاصیت محلی
 global property خاصیت سراسری
 self-adaptive..... خودتنظیمی
 relative error..... خطای نسبی

ر

Spectral conjugate gradient method..... روش گرادیان مزدوج طیفی
 nonlinear conjugate gradient method..... روش گرادیان مزدوج غیرخطی

hybrid conjugate gradient method روش گرادیان مزدوج هیبریدی
 structure spectrul method روش طیفی ساختاریافته
 image deblurring رفع تاری تصویر

ز

CPU time زمان پردازش

س

Newton structure ساختار نیوتنی

ش

Generalized conjugacy condition شرط همگرایی تعمیم‌یافته
 conjugacy condition شرط هم‌ارزی
 standard Wolfe condition شرط استاندارد ولف
 curvature condition شرط انحنا
 string Wolfe condition شرط ولف قوی
 Neural Network شبکه عصبی

ع

computational performance عملکرد محاسباتی

ف

adaptive median filter فیلتر میانگین تطبیقی

ک

conjugate descent کاهش مزدوج
 Alternating-reduction کاهش تناوبی

minimizer کمینه کننده
 local minimizer کمینه‌کننده محلی
 strongly convex کوژ شدید
 robotic motion control کنترل حرکت رباتیک
 minimum of noisy pixel کمترین مقدار نویز پیکسلی

گ

CG-DESCENT گرادیان مزدوج کاهشی

م

large scale مقیاس بزرگ
 secant equation معادله سکانتی
 level set مجموعه سطح

ن

monotonically decreasing نزولی یکنواخت
 noise condidate نویز کاندید
 peak signal-to-noise ratio نسبت اوج سیگنال به نویز

و

double-truncated property ویژگی دوگانه برش یافته
 strong global convergence property ویژگی قوی همگرایی سراسری
 robust convergence property ویژگی همگرایی پایدار

ه

Global convergence همگرایی سراسری