

See discussions, stats, and author profiles for this publication at:  
<http://www.researchgate.net/publication/260752837>

# A modified Polak–Ribière–Polyak conjugate gradient algorithm for nonsmooth convex programs

ARTICLE *in* JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS · FEBRUARY 2014

Impact Factor: 1.27 · DOI: 10.1016/j.cam.2013.04.032

CITATIONS

12

READS

97

3 AUTHORS, INCLUDING:



Zengxin Wei

Guangxi University

74 PUBLICATIONS 856 CITATIONS

SEE PROFILE



G. Y. Li

University of New South Wales

79 PUBLICATIONS 744 CITATIONS

SEE PROFILE

All in-text references [underlined in blue](#) are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: Gonglin Yuan  
Retrieved on: 28 December 2015

# یک الگوریتم شیب مزدوج Polak-Ribere-Polyak اصلاح شده برای برنامه های محدب غیر صاف

Gonglin Yuan<sup>a</sup>, Zengxin Wei<sup>a,1</sup>, Guoyin Li<sup>b,\*</sup>

<sup>a</sup> Department of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, 530004, PR China

<sup>b</sup> Department of Applied Mathematics, The University of New South Wales, Kensington, 2052, Sydney, Australia

## چکیده

روش گرادیان مزدوج (CG)، یکی از محبوب ترین روش ها برای حل مسائل بهینه سازی نامحدود صاف، به دلیل سادگی و نیاز به حافظه کم آن است.

با این حال، استفاده از روش های CG، تا اینجا عمدتاً به حل مسائل بهینه سازی صاف محدود شده است. هدف این مقاله، ارائه روش های شیب-نوع مزدوج کارآمد برای حل مسائل بهینه سازی غیر صاف است. با استفاده از رویکرد مقررات مورو-Yosida (صاف) و یک تکنیک جستجو خط غیر صاف، ما یک الگوریتم گرادیان مزدوج Polak-Ribière-پولیاک (PRP) اصلاح شده برای حل یک مشکل حداقل سازی محدب نامحدود غیر صاف ارائه کردیم. الگوریتم ما دارای سه ویژگی مطلوب است. (۱) روش جستجو، ویژگی نزول کافی را برآورده می کند و به یک منطقه اعتماد به صورت خود کار تعلق دارد؛ (۲) روش جستجو، نه تنها از اطلاعات گرادیان استفاده می کند، بلکه همچنین از اطلاعات مقدار تابع نیز استفاده می کند؛ و (۳) الگوریتم، یک ویژگی مهم از روش PRP به خوبی شناخته شده را به همراه دارد: گرایش به سمت شیب دار ترین جهت نزولی است، اگر یک گام کوچک به دور از راه حل ایجاد شود، که از رخ دادن توالی از مراحل کوچک جلوگیری می کند. تحت شرایط استاندارد، ما نشان دادیم که الگوریتم بصورت کلی به یک راه حل بهینه همگرا است. آزمایش عددی نشان می دهد که الگوریتم ما برای حل مسائل بهینه سازی محدب نامحدود غیر صاف مقیاس بزرگ موثر و مناسب است.

## مقدمه

مسئله بهینه سازی محدب نامحدود زیر را در نظر بگیرید:

$$(P) \quad \min_{x \in \mathbb{R}^n} f(x), \quad (1.1)$$

که در آن  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  یک تابع محدب غیر صاف محتمل است. در موارد خاص زمانی که  $f$  پیوسته مشتق پذیر است، این مسئله بهینه سازی به خوبی برای چندین دهه مورد مطالعه قرار گرفته است. به طور خاص، روش گرادیان مزدوج (CG) و روش شبه نیوتن دو روش عمده محبوب برای حل مسائل بهینه سازی نامحدود محدب

صاف هستند. روش شبه نیوتن شامل محاسبه / تقریبی از ماتریس هشین از تابع هدف است و اغلب دارای همگرایی سریع است. از سوی دیگر، روش گرادیان مزدوج تنها از اطلاعات دسته اول استفاده می کند و بنابراین برای حل مسائل بهینه سازی در مقیاس بزرگ مناسب است. در حال حاضر، بسیاری از فرمول های گرادیان مزدوج خوب-شناخته شده ای وجود دارند (برای مثال، نگاه کنید به [1-10]). به طور خاص، روش گرادیان مزدوج به اصطلاح پولاک-ریبری-پولیاک (PRP) شده است به خوبی مورد مطالعه قرار گرفته و به طور کلی معتقد به داشتن یکی از کارآمدترین روش های گرادیان مزدوج است. با این حال، استفاده از روش های گرادیان مزدوج تاکنون عمدتاً به حل مسائل بهینه سازی صاف محدود شده است.

به تازگی، بسیاری از برنامه های پیشرفته بهینه سازی برای نیاز به مطالعه مساله بهینه سازی محدب غیر صاف مقیاس-بزرگ نامیده می شود. به عنوان یک مثال توضیحی، اجازه دهید مساله بازسازی تصویر ناشی در پردازش تصویر را در نظر بگیریم. مساله بازسازی تصویر، جهت بازسازی یک تصویر از یک صحنه ناشناخته بدست آمده از یک تصویر مشاهده شده است. این مساله، نقش مهمی را در علوم پزشکی، مهندسی زیستی و زمینه های دیگری از علوم و مهندسی ایفا می کند [۱۱-۱۳]. رایج ترین تصویر الگوی تنزل می تواند توسط سیستم زیر نشان داده شود:

$$b = Ax + \eta$$

که در آن  $\eta \in \mathbb{R}^m$  نشان دهنده نویز،  $A$  یک ماتریس  $n \times m$  مبهم است،  $x \in \mathbb{R}^n$  و  $b \in \mathbb{R}^m$  به ترتیب تصاویر مشاهده شده و اصلی هستند. همانطور که  $\eta$  معمولاً ناشناخته است، یک راه برای حل آن، حل مسئله بهینه سازی مربع  $\min_{x \in \mathbb{R}^n} \|Ax + b\|^2$  است. حل این مساله به تنهایی یک راه حل رضایت بخش را ارائه نمی دهد، چونکه سیستم به سر و صدا و عدم وجود اطلاعات بسیار حساس است. برای غلبه بر این، مساله مربع حداقل تنظیم شده زیر اغلب استفاده می شود:

$$\min_{x \in \mathbb{R}^n} \|Ax + b\|^2 + \lambda \|Dx\|_1$$

که در آن  $D$  یک عملگر خطی است،  $\lambda$  شاخص تنظیم است که مبادله بین اصطلاح داده های مناسب و اصطلاح تنظیم را کنترل می کند و  $\|\cdot\|_1$ ،  $l^1$  معمول است. همانطور که  $l^1$  معمول غیر صاف است، مساله بالا یک مسئله بهینه سازی محدب غیر صاف است و معمولاً مقیاس بزرگ است.

هدف از این مقاله ارائه روش های نوع گرادیان مزدوج کارآمد برای حل مساله بهینه سازی غیر صاف (P) است. با استفاده از رویکرد تنظیم مورو- Yosida (صاف) و یک تکنیک جستجو خط غیر صاف، ما یک الگوریتم گرادیان مزدوج PRP برای حل یک مساله حداقل سازی محدب نامحدود غیر صاف را ارائه کردیم.

الگوریتم ما دارای سه ویژگی مورد نظر زیر است. (۱) جهت جستجو، ویژگی به اندازه کافی تنزلی را برآورده می کند و متعلق به یک منطقه اعتماد به صورت خودکار است؛ (۲) جهت جستجو، استفاده از نه تنها اطلاعات گرادیان را ممکن می سازد بلکه از اطلاعات تابع نیز بهره می برد؛ و (۳) این الگوریتم، یک ویژگی مهم از روش PRP به خوبی شناخته شده را به همراه دارد:

گرایش به سمت شدیدترین جهت تنزل می چرخد، اگر یک گام کوچک به دور از راه حل ایجاد شود، که از رخ دادن توالی مراحل کوچک جلوگیری می کند.

این مقاله به شرح زیر است. در بخش ۲، ما به طور خلاصه برخی از نتایج اساسی در تجزیه و تحلیل محدب و تجزیه و تحلیل غیر صاف را بررسی می کنیم. در بخش ۳، ما الگوریتم جدید گرادیان مزدوج PRP اصلاح شده را ارائه می کنیم. در بخش ۴، ما همگرایی کلی روش ارائه شده را اثبات می کنیم. در بخش ۵، ما برخی از ضمیمه های مشابه را مورد بحث قرار می دهیم و سه فرمول گرادیان مزدوج اصلاح شده دیگر را ارائه می کنیم. در بخش ۶، ما نتایج عددی را برای الگوریتم مان گزارش می کنیم و چند مقایسه برای روش های موجود برای هر دو مقیاس کوچک و بزرگ مسائل بهینه سازی محدب غیر صاف ارائه می کنیم. در نهایت، ما مقاله مان را نتیجه گیری می کنیم و برخی از موضوعات تحقیقاتی ممکن در آینده را در بخش ۷ ذکر می کنیم. سراسر این مقاله، بدون مشخصات،  $\| \cdot \|$  نشان دهنده بردارها یا ماتریس های اقلیدس معمول است.

## 2. عناصر تجزیه و تحلیل محدب و تجزیه و تحلیل غیر صاف

در این بخش، ما برخی از حقایق و نتایج اساسی در تجزیه و تحلیل محدب و تجزیه و تحلیل غیر صاف را بررسی می کنیم. اجازه دهید  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  یک تابع محدب باشد (احتمالاً غیر صاف). اجازه دهید  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  تنظیم به اصطلاح مورو - Yosida تعریف شده از  $f$  توسط

$$F(x) = \min_{z \in \mathbb{R}^n} \left\{ f(z) + \frac{1}{2\lambda} \|z - x\|^2 \right\},$$

باشد که در آن  $\lambda$  یک شاخص مثبت و  $\| \cdot \|$  اقلیدس معمول را نشان می دهد. اجازه دهید

$$\theta(z, x) = f(z) + \frac{1}{2\lambda} \|z - x\|^2$$

و نشان می دهد  $p(x) = \operatorname{argmin}_z \theta(z, x)$ . آنگاه،  $p(x)$  به خوبی تعریف شده و منحصر به فرد است، از آنجا که  $\theta$  برای هر  $x$  ثابت بشدت محدب است.

با معادله (2.1)،  $F$  را می توان بصورت

$$F(x) = f(p(x)) + \frac{1}{2\lambda} \|p(x) - x\|^2.$$

بیان کرد.

در ادامه، ما شیب  $F$  را توسط  $g$  نشان می دهیم. برخی از خواص مهم و مفید تابع  $F$  تنظیم مورو-Yosida به صورت زیر تعیین شده است.

(۱) تابع  $F$  مقدار محدود، محدب، و در همه جا مشتق پذیر است با

$$g(x) = \nabla F(x) = \frac{x - p(x)}{\lambda}. \quad (2.2)$$

علاوه بر این، ترسیم شیب  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  بصورت کلی Lipschitz مستمر با مدول  $\lambda$  است، یعنی،

$$\|g(x) - g(y)\| \leq \frac{1}{\lambda} \|x - y\|, \quad \forall x, y \in \mathbb{R}^n. \quad (2.3)$$

(۲)  $x$ ،  $\min_{x \in \mathbb{R}^n} f(x)$  را حل می کند، اگر و تنها اگر  $\nabla F(x) = 0$ ، یعنی  $p(x) = x$ . واضح است که  $F(x)$  و  $g(x)$  را می توان از طریق راه حل بهینه  $\operatorname{argmin}_{z \in \mathbb{R}^n} \theta(z)$  بدست آورد. با این حال،  $p(x)$  حداقل کننده  $\theta(z)$ ، حل دقیق دشوار یا حتی غیر ممکن است. بنابراین، در محاسبات واقعی، به جای محاسبه مقدار دقیق  $p(x)$ ،  $F(x)$  و  $g(x)$  ما اغلب از برخی تقریب های متناسب استفاده می کنیم. در واقع، برای هر  $x \in \mathbb{R}^n$  و هر  $\varepsilon > 0$ ، در اینجا یک بردار  $p^\alpha(x, \varepsilon) \in \mathbb{R}^n$  وجود دارد بطوری که

$$f(p^\alpha(x, \varepsilon)) + \frac{1}{2\lambda} \|p^\alpha(x, \varepsilon) - x\|^2 \leq F(x) + \varepsilon. \quad (2.4)$$

بنابراین، هنگامی که  $\varepsilon$  کوچک است، ما می توانیم از  $p^\alpha(x, \varepsilon)$  برای تعریف تقریب های  $F(x)$  و  $g(x)$  بترتیب بشرح زیر استفاده کنیم:

$$F^\alpha(x, \varepsilon) = f(p^\alpha(x, \varepsilon)) + \frac{1}{2\lambda} \|p^\alpha(x, \varepsilon) - x\|^2 \quad (2.5)$$

و

$$g^\alpha(x, \varepsilon) = \frac{x - p^\alpha(x, \varepsilon)}{\lambda}, \quad (2.6)$$

برخی از الگوریتم های قابل پیاده سازی برای محاسبه  $p^\alpha(x, \varepsilon)$  برای یک تابع محدب غیر قابل مشتق پذیر را می توان در [۱۴] یافت. یک ویژگی قابل توجه  $F^\alpha(x, \varepsilon)$  و  $g^\alpha(x, \varepsilon)$  به شرح زیر ارائه شده است [15].

**تناسب ۱.۲.** اجازه دهید  $p^\alpha(x, \varepsilon)$  یک بردار برآورده کننده (2.4) باشد و  $F^\alpha(x, \varepsilon)$  و  $g^\alpha(x, \varepsilon)$  بترتیب توسط

(۲,۵) و (۲,۶) تعریف شده اند.

سپس ما داریم

$$F(x) \leq F^\alpha(x, \varepsilon) \leq F(x) + \varepsilon, \quad (2.7)$$

$$\|p^\alpha(x, \varepsilon) - p(x)\| \leq \sqrt{2\lambda\varepsilon}, \quad (2.8)$$

و

$$\|g^\alpha(x, \varepsilon) - g(x)\| \leq \sqrt{2\varepsilon/\lambda}. \quad (2.9)$$

تناسب فوق نشان می دهد که ما می توانیم تقریب های  $F^\alpha(x, \varepsilon)$  و  $g^\alpha(x, \varepsilon)$  را محاسبه کنیم. با انتخاب پارامتر  $\varepsilon$  به اندازه کافی کوچک  $F^\alpha(x, \varepsilon)$  و  $g^\alpha(x, \varepsilon)$  ممکن است به صورت خودسرانه به  $F(x)$  و  $g(x)$  نزدیک باشند. بر اساس این ویژگی ها، الگوریتم های بسیاری برای حل مساله (P) ارائه شده است (برای مثال، نگاه کنید به [16]). الگوریتم نقطه نزدیک کلاسیک [۱۷] می تواند به عنوان یک روش نوع گرادیان برای حل مساله (P) در نظر گرفته شود و ثابت شود که در برخورد با مشکل ارزیابی مقدار تابع  $F(x)$  و شیب آن  $\nabla F(x)$  در یک نقطه معین  $x$  موثر می باشد (نگاه کنید به [18-22]). لمار کال [۲۳] و ولف [۲۴] یک گام رو به جلو بسیار بزرگ را در بهینه سازی غیر صاف بوسیله مفهوم دسته آغاز کردند، که می تواند  $F$  محدب و غیر محدب را مدیریت کند. علاوه بر این، کیویل [۲۵] یک متغیر دسته را ارائه کرد، که نزدیک به روش تکرار اعتماد دسته است (نگاه کنید به [26]). به تازگی، Haarala و همکاران، یک روش دسته حافظه محدود جدید را برای مسائل بهینه سازی غیر صاف مقیاس بزرگ پیشنهاد کردند [۲۷]. برای مراجع دیگر روش های دسته، شخص ممکن است به [۲۸-۳۰] مراجعه کند.

### 3. روش PRP اصلاح شده برای مساله غیر صاف

در این بخش، ما روش های گرادیان مزدوج PRP اصلاح شده را برای حل مسئله بهینه سازی محدب غیر صاف نامحدود  $\min_{x \in \mathbb{R}^n} f(x)$  معرفی می کنیم، که در آن  $f$  یک تابع محدب است. ایده اصلی ما امکان استفاده از تنظیم مورو-Yosida برای صاف کردن تابع و همچنین دو روش شتاب به تازگی معرفی شده در اصلاح روش های CG است. برای این کار، ما نخست دو روش شتاب برای اصلاح روش های CG در حل یک مسئله بهینه سازی صاف معرفی می کنیم.

۳.۱. روش اصلاح شده CG با اعمال شرط تنزل کافی

به یاد بیاورید که وقتی تابع  $f$  صاف است، جهت جستجو روش های گرادیان مزدوج PRP معروف است با

$$d_{k+1} = \begin{cases} -\nabla f_{k+1} + \beta_k^{\text{PRP}} d_k, & \text{if } k \geq 1, \\ -\nabla f_{k+1}, & \text{if } k = 0, \end{cases} \quad (3.1)$$

تعیین می شود

که در آن  $\nabla f_k = \nabla f(x_k)$  and  $\beta_k^{\text{PRP}} = \frac{\nabla f_{k+1}^T y_k}{\|\nabla f_k\|^2}$  and  $y_k = \nabla f_{k+1} - \nabla f_k$ . به طور کلی، روش PRP ممکن نیست بصورت کلی همگرا باشد زمانیکه آن شرط به اصطلاح کافی زیر را برآورده نمی کند: در اینجا  $r > 0$  وجود دارد به طوری که

$$\nabla f_k^T d_k \leq -r \|\nabla f_k\|^2 \quad \text{for all } k \in \mathbb{N}.$$

توجه داشته باشید که شرط تنزل به اندازه کافی معمولاً نقش مهمی را در تجزیه و تحلیل همگرا کلی روش های گرادیان مزدوج بازی می کند، و همچنین بسیاری از نویسندگان اشاره کردند که شرط تنزل به اندازه کافی ممکن است برای روش های CG بسیار مهم باشد [31-34]. به منظور اطمینان حاصل شدن از این که جهت جستجو این ویژگی را دارد، بسیاری از روش های CG اصلاح شده ارائه شده اند ([35-41] و غیره)، که در آن ژانگ و همکاران [41] یک روش PRP سه دوره ای را ارائه کردند که در آن جهت جستجو توسط

$$d_{k+1} = \begin{cases} -\nabla f_{k+1} + \beta_k^{\text{PRP}} d_k - \vartheta_k y_k, & \text{if } k \geq 1 \\ -\nabla f_{k+1}, & \text{if } k = 0, \end{cases} \quad (3.2)$$

تعریف شده است، که در آن  $\vartheta_k = \frac{\nabla f_{k+1}^T d_k}{\|\nabla f_k\|^2}$ ,  $v_{k+1} = v_k + \mu_k d_k$ , and  $\mu_k$  گام اندازه است. می توان تایید کرد که برای همه  $k$  ها  $d_k^T \nabla f_k = -\|\nabla f_k\|^2$ . این روش، روش PRP استاندارد را اگر جستجوی خط دقیق استفاده شود، کاهش می دهد؛ همگرایی کلی اش با جستجو خط نوع Armijo در [41] به دست آمده است.

### 2.3. روش اصلاح شده CG با ترکیب اطلاعات مقدمات تابع

قبل از اینکه ما به الگوریتم PRP اصلاح شده مان ادامه دهیم، اجازه دهید تکنیک تسریع دوم در اصلاح روش گرادیان مزدوج که توسط نویسندگان در [42] استفاده شده است، را معرفی کنیم. به یاد بیاورید که روش موثر دیگر برای حل یک مسئله بهینه سازی صاف، روش وتری شبه نیوتن است که در آن تکرار  $x_{k+1} = x_k - \alpha_k$ ،  $B_k^{-1} \nabla f_k$  را برآورده می کند، که در آن  $B_k$ ، که یک تقریب هسی از  $f$  در  $x_k$  است. توالی ماتریس  $\{B_k\}$  معادله متقاطع زیر را برآورده می کند:

$$B_{k+1} s_k = y_k, \quad (3.3)$$

که در آن  $s_k = x_{k+1} - x_k$  است. بدیهی است، تنها اطلاعات گرادیان در  $(3,3)$  بهره برداری شده است، در حالی که مقادیر تابع در دسترس، در نظر گرفته نشده اند. از این رو، روش های استفاده کننده از مقادیر گرادیان و همچنین مقادیر تابع توسط چندین محقق مطالعه شده است. یک تلاش قابل توجه که معادله قاطع معمول را با استفاده از هر دو مقادیر تابع و مقادیر شیب اصلاح کرده است، بخاطر وی و همکارانش است (نگاه کنید به [43])، که در آن معادله قاطع توسط

$$B_{k+1} s_k = y_k^*, \quad (3.4)$$

تعریف شده است

که در آن  $y_k^* = y_k + \gamma_k^* s_k$  and  $\gamma_k^* = \frac{(\nabla f(x_{k+1}) + \nabla f(x_k))^T s_k + 2(f(x_k) - f(x_{k+1}))}{\|s_k\|^2}$ . یک ویژگی قابل توجه از این معادله قاطع (۳,۴) این است که اگر  $f$  دو بار پیوسته مشتق پذیر باشد و  $B_{k+1}$  توسط فرمول روش BFGS به روز شود، سپس تساوی زیر

$$f(x_k) = f(x_{k+1}) + \nabla f(x_{k+1})^T s_k + \frac{1}{2} s_k^T B_{k+1} s_k \quad (3.5)$$

برای همه  $k$  ها برقرار است. علاوه بر این، این ویژگی مستقل از هر فرض تحذب در تابع هدف است. مزیت نظری معادله جدید شبه نیوتن (۳,۴) را می توان از قضیه زیر دید.

**قضیه ۱.۳** ([43]) فرض کنید که تابع  $f$  به اندازه کافی صاف است و  $\|s_k\|$  به اندازه کافی کوچک است، پس ما داریم

$$s_k^T \nabla^2 f_{k+1} s_k - s_k^T y_k^* - \frac{1}{3} s_k^T (T_{k+1} s_k) s_k = O(\|s_k\|^4) \quad (3.6)$$

و

$$s_k^T \nabla^2 f_{k+1} s_k - s_k^T y_k - \frac{1}{2} s_k^T (T_{k+1} s_k) s_k = O(\|s_k\|^4) \quad (3.7)$$

که در آن  $\nabla^2 f_{k+1}$  ماتریس هسی  $f$  در  $x_{k+1}$ ،  $T_{k+1}$  را نشان می دهد، تانسور  $f$  در  $x_{k+1}$  و

$$s_k^T (T_{k+1} s_k) s_k = \sum_{i,j,l=1}^n \frac{\partial^3 f(x_{k+1})}{\partial x^i \partial x^j \partial x^l} s_k^i s_k^j s_k^l.$$

نتیجه بالا نشان می دهد که معادله جدید شبه نیوتن (۳,۴) رابطه تقریبی بهتری نسبت به (۳,۳) دارد. بر اساس این معادله جدید شبه نیوتن، روش های مختلف کارآمد جدیدی از گرادیان مزدوج پیشنهاد شدند و بوسیله جایگزین کردن  $y_k$  با  $y_k^*$  به بهبود عددی منجر شده اند (نگاه کنید به [۴۴-۴۷] و غیره). تحریک بوسیله فرمول PRP سه دوره ای (۳,۲) و معادله شبه نیوتن جدید (۳,۴)، ما اکنون یک روش PRP سه دوره ای اصلاح شده را برای حل یک مسئله بهینه سازی غیر صاف ارائه می کنیم.

### 3.3. روش PRP جدید برای یک برنامه محذب غیر صاف

به یاد بیاورید که  $g^\alpha(x_k, \varepsilon_k)$  یک تقریب از شیب  $F$  در  $x_k$  است (تنظیم مورو Yosida- از تابع هدف  $f$ ). با جایگزین کردن  $\nabla f_k$  و  $y_k$  بترتیب با  $g^\alpha(x_k, \varepsilon_k)$  و  $y_k^*$  در فرمول PRP سه دوره ای (۳,۲)، ما اکنون یک فرمول گرادیان مزدوج PRP اصلاح شده را برای حل (P) بشرح زیر ارائه می کنیم:



$$d_{k+1} = \begin{cases} -g^\alpha(x_{k+1}, \varepsilon_{k+1}) + \frac{g^\alpha(x_{k+1}, \varepsilon_{k+1})^T y_k^* d_k - d_k^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) y_k^*}{\max\{2c \|d_k\| \|y_k^*\|, \|g^\alpha(x_k, \varepsilon_k)\|^2\}}, & \text{if } k \geq 1 \\ -g^\alpha(x_{k+1}, \varepsilon_{k+1}), & \text{if } k = 0, \end{cases} \quad (3.8)$$

که در آن  $d_k = x_{k+1} - x_k$ ,  $s_k = g^\alpha(x_{k+1}, \varepsilon_{k+1}) - g^\alpha(x_k, \varepsilon_k)$ ,  $y_k = g^\alpha(x_{k+1}, \varepsilon_{k+1}) - g^\alpha(x_k, \varepsilon_k)$ ,  $y_k^* = y_k + \gamma_k^* s_k$ , جهت تحقیق در  $k$  مین،  $c > 0$  یک ثابت است، و

$$\gamma_k^* = \frac{(g^\alpha(x_{k+1}, \varepsilon_{k+1}) + g^\alpha(x_k, \varepsilon_k))^T s_k + 2(F^\alpha(x_k, \varepsilon_k) - F^\alpha(x_{k+1}, \varepsilon_{k+1}))}{\|s_k\|^2}.$$

شایان ذکر است که اصطلاح سنجش  $\|\nabla f(x_k)\|^2$  در مخرج فرمول PRP سه دوره ای (۳،۲) با  $\max\{2c \|d_k\| \|y_k^*\|, \|g^\alpha(x_k, \varepsilon_k)\|^2\}$  در فرمول گرادیان مزدوج PRP جدیدمان سازگار شده است. این اصلاح جهت نشان دادن اینکه همه جهت های جستجو در یک منطقه اعتماد به طور خودکار باقی خواهد ماند، به ما کمک خواهد کرد (گزاره ۳،۱ را ببینید).

این الگوریتم از روش گرادیان مزدوج PRP اصلاح شده به شرح زیر بیان شده است.

**الگوریتم ۱.** الگوریتم گرادیان مزدوج غیر یکنواخت.

گام ۰. انتخاب کنید  $x_0 \in \mathbb{R}^n$ ,  $\varepsilon_0 \in (0, 1)$ ,  $\sigma \in (0, 1)$ ,  $c > 0$ ,  $s > 0$ ,  $\lambda > 0$ ,  $\rho \in [0, 1]$ ,  $E_0 = 1$ ,  $J_0 = F^\alpha(x_0, \varepsilon_0)$ ,  $d_0 = -g^\alpha(x_0, \varepsilon_0)$  و  $\mathcal{E} \in (0, 1)$  اجازه دهید  $k=0$ .

گام ۱. اگر  $\|g^\alpha(x_k, \varepsilon_k)\| < \varepsilon$  شرط خاتمه  $\mathcal{E}$  برآورده کند، سپس متوقف شود. در غیر این صورت، به گام ۲ بروید.  
گام ۲. یک کمیت  $\mathcal{E}_{k+1}$  برآورده کننده  $0 < \varepsilon_{k+1} < \varepsilon_k$  را انتخاب کنید، و اندازه گام  $\alpha_k$  را بوسیله جستجو خط نوع Armijo غیر یکنواخت زیر، محاسبه کنید:

$$F^\alpha(x_k + \alpha_k d_k, \varepsilon_{k+1}) - J_k \leq \sigma \alpha_k g^\alpha(x_k, \varepsilon_k)^T d_k, \quad (3.9)$$

که در آن  $\alpha_k = s 2^{-i_k}$ ,  $i_k \in \{0, 1, 2, \dots\}$ .

گام ۳. اجازه دهید  $x_{k+1} = x_k + \alpha_k d_k$ . If  $\|g^\alpha(x_{k+1}, \varepsilon_{k+1})\| < \varepsilon$ ، سپس متوقف شود. در غیر این صورت، به گام ۴ بروید.

گام ۴.  $J_k$  را با استفاده از فرمول زیر بروز رسانی کنید:

$$E_{k+1} = \rho E_k + 1, \quad J_{k+1} = \frac{\rho E_k J_k + F^\alpha(x_k + \alpha_k d_k, \varepsilon_{k+1})}{E_{k+1}}. \quad (3.10)$$

گام ۵. جهت جستجو  $d_{k+1}$  را توسط (۳،۸) محاسبه کنید.

گام ۶. اجازه دهید  $K := K + 1$  و به مرحله ۱ برگردید.

**تبصره ۱.** روش جستجو خط (۳،۹) توسط ژانگ و هاگر [۴۸] تحریک شده است. دیدن اینکه  $J_{k+1}$  یک ترکیب محدب  $J_k$  و  $F^\alpha(x_{k+1}, \varepsilon_{k+1})$  است، دشوار می باشد. با توجه به  $J_0 = F^\alpha(x_0, \varepsilon_0)$ ، متعاقب آن  $J_k$  یک ترکیب محدب از

مقادیر تابع  $F^\alpha(x_0, \varepsilon_0), F^\alpha(x_1, \varepsilon_1), \dots, F^\alpha(x_k, \varepsilon_k)$  است. انتخاب  $\rho$ ، درجه غیر یکنواختی را کنترل می کند. اگر  $\rho = 0$ ، آنگاه جستجو خط، جستجو خط یکنواخت Armijo معمول است. اگر  $\rho = 1$ ، آنگاه  $J_k = C_k$ ، که در آن

$$C_k = \frac{1}{k+1} \sum_{i=0}^k F^\alpha(x_i, \varepsilon_i)$$

مقدار میانگین تابع است.

گزاره زیر نشان می دهد که روش CG ما، خاصیت تنزل به اندازه کافی را برآورده می کند و جهت جستجوی CG متناظر، متعلق به یک منطقه اعتماد به صورت خود کار است.

گزاره ۳,۱ (ویژگی تنزل کافی). برای تمام  $k \in \mathbb{N} \cup \{0\}$ ، ما داریم

$$g^\alpha(x_k, \varepsilon_k)^T d_k = -\|g^\alpha(x_k, \varepsilon_k)\|^2 \quad (3.11)$$

و

$$\|d_k\| \leq \left(1 + \frac{1}{c}\right) \|g^\alpha(x_k, \varepsilon_k)\|. \quad (3.12)$$

**اثبات.** برای  $k = 0$ ، ما داریم  $d_0 = -g^\alpha(x_0, \varepsilon_0)$ . بنابراین، (۳,۱۱) و (۳,۱۲) به وضوح برقرار می شود. برای  $k \geq 1$ ،

با تعریف  $d_k$ ، ما بدست می آوریم

$$\begin{aligned} d_{k+1}^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) &= -\|g^\alpha(x_{k+1}, \varepsilon_{k+1})\|^2 + \left[ \frac{g^\alpha(x_{k+1}, \varepsilon_{k+1})^T y_k^* d_k - d_k^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) y_k^*}{\max\{2c \|d_k\| \|y_k^*\|, \|g^\alpha(x_k, \varepsilon_k)\|^2\}} \right]^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) \\ &= -\|g^\alpha(x_{k+1}, \varepsilon_{k+1})\|^2. \end{aligned} \quad (3.13)$$

بنابراین (۳,۱۱) برقرار می شود. اکنون ما (۳,۱۲) را نشان می دهیم. با استفاده از تعریف  $d_k$  دوباره، ما داریم

$$\begin{aligned} \|d_{k+1}\| &= \left\| -g^\alpha(x_{k+1}, \varepsilon_{k+1}) + \frac{g^\alpha(x_{k+1}, \varepsilon_{k+1})^T y_k^* d_k - d_k^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) y_k^*}{\max\{2c \|d_k\| \|y_k^*\|, \|g^\alpha(x_k, \varepsilon_k)\|^2\}} \right\| \\ &\leq \|g^\alpha(x_{k+1}, \varepsilon_{k+1})\| + \frac{\|g^\alpha(x_{k+1}, \varepsilon_{k+1})\| \|y_k^*\| \|d_k\| + \|d_k\| \|g^\alpha(x_{k+1}, \varepsilon_{k+1})\| \|y_k^*\|}{\max\{2c \|d_k\| \|y_k^*\|, \|g^\alpha(x_k, \varepsilon_k)\|^2\}} \\ &\leq \left(1 + \frac{1}{c}\right) \|g^\alpha(x_{k+1}, \varepsilon_{k+1})\|, \end{aligned}$$

که در آن آخرین نابرابری بشرح زیر است

$$\max\{2c \|d_k\| \|y_k^*\|, \|g^\alpha(x_k, \varepsilon_k)\|^2\} \geq 2c \|d_k\| \|y_k^*\|.$$

اثبات، کامل است.

#### ۴. تجزیه و تحلیل همگرایی

در این بخش، ما یک تجزیه و تحلیل همگرایی جامع را برای الگوریتم اصلاح شده PRP مان ارائه می کنیم. به منظور رسیدن به همگرایی کلی الگوریتم ۱، فرض زیر از سراسر این بخش ساخته شده است.

**فرض A.** (الف) در اینجا یک ثابت مثبت  $M$  وجود دارد به طوری که

$$\|\nabla F(u_k)\| \leq M, \quad \forall u_k \in [x_k, x_{k+1}] \text{ and } \forall k \in \mathbb{N} \quad (4.1)$$

که در آن  $F$  تنظیم مورو-یوسیدا  $f$  است.

(ب)  $F$  از زیر محدود شده است.

(ج) دنباله  $\varepsilon_k$  به صفر همگرا می شود.

**تبصره دوم.** اگر  $s_k \rightarrow 0$  دشوار است که بدست آوریم

$$\|g^\alpha(x_{k+1}, \varepsilon_{k+1})\| - \|g^\alpha(x_k, \varepsilon_k)\| \rightarrow 0 \text{ and } F^\alpha(x_{k+1}, \varepsilon_{k+1}) - F^\alpha(x_k, \varepsilon_k) \rightarrow 0.$$

آنگاه، ما می بینیم که

$$\|d_{k+1} - (-g^\alpha(x_{k+1}, \varepsilon_{k+1}))\| \rightarrow 0.$$

این نشان می دهد که روش ارائه شده، ویژگی روش گرادیان مزدوج PRP بخوبی شناخته شده را به همراه دارد: گرایش به سمت شدیدترین جهت تنزل می چرخد، اگر یک گام کوچک به دور از راه حل ایجاد شود، که از رخ دادن یک توالی از مراحل کوچک جلوگیری می کند.

با استفاده از گزاره ۳،۱ و فرض A، مشابه گزاره ۱،۱ در [۴۸]، ما می توانیم نشان دهیم که الگوریتم ۱ بعنوان به خوبی تعریف شده در گزاره زیر است. همانطور که این اثبات، در اصل همان گزاره ۱،۱ در [۴۸] است، ما اثباتش را در اینجا حذف کردیم.

**گزاره ۴،۱.** فرض کنید که فرض A برقرار است. سپس، برای تکرار های ایجاد شده توسط الگوریتم ۱، ما برای هر یک از  $k$  ها داریم  $F^\alpha(x_k, \varepsilon_k) \leq J_k \leq C_k$ ، که در آن  $C_k = \frac{1}{k+1} \sum_{i=0}^k F^\alpha(x_i, \varepsilon_i)$ . علاوه بر این، در اینجا  $\alpha_k$  برآورده کننده شرایط Armijo از بروز رسانی جستجو خط وجود دارد.

**گزاره ۴،۲.** فرض کنید که فرض A برقرار است. اجازه دهید  $\{(x_k, \alpha_k)\}$  دنباله ایجاد شده توسط الگوریتم ۱ باشد. فرض کنید که  $\varepsilon_k = o(\alpha_k^2 \|d_k\|^2)$  برقرار است. آنگاه در اینجا یک ثابت  $m_0 > 0$  وجود دارد بطوری که

$$\alpha_k \geq m_0$$

**اثبات.** اجازه دهید  $\alpha_k$  جستجو خط (۳،۹) را برآورده کند. ما بوسیله روش تناقض پیش می رویم و فرض می کنیم

که  $\liminf_{k \rightarrow \infty} \alpha_k = 0$ .

با عبور بسمت یک توالی در صورت لزوم، ممکن است فرض کنیم که  $\alpha_k \rightarrow 0$ . آنگاه، بوسیله جستجو خط،  $\alpha'_k = \frac{\alpha_k}{2}$  بر آورده می کند

$$F^\alpha(x_k + \alpha'_k d_k, \varepsilon_{k+1}) - J_k > \sigma \alpha'_k g^\alpha(x_k, \varepsilon_k)^T d_k.$$

با استفاده از  $F^\alpha(x_k, \varepsilon_k) \leq J_k \leq C_k$  در گزاره ۴,۱ ما بدست می آوریم

$$F^\alpha(x_k + \alpha'_k d_k, \varepsilon_{k+1}) - F^\alpha(x_k, \varepsilon_k) \geq F^\alpha(x_k + \alpha'_k d_k, \varepsilon_{k+1}) - J_k > \sigma \alpha'_k g^\alpha(x_k, \varepsilon_k)^T d_k. \quad (4.3)$$

بوسیله (۴,۳)، (۲,۷) و فرمول تیلور، ما داریم

$$\begin{aligned} \sigma \alpha'_k g^\alpha(x_k, \varepsilon_k)^T d_k &< F^\alpha(x_k + \alpha'_k d_k, \varepsilon_{k+1}) - F^\alpha(x_k, \varepsilon_k) \\ &\leq F(x_k + \alpha'_k d_k) - F(x_k) + \varepsilon_{k+1} \\ &= \alpha'_k d_k^T g(x_k) + \frac{1}{2} (\alpha'_k)^2 d_k^T \nabla F(u_k) d_k + \varepsilon_{k+1} \\ &\leq \alpha'_k d_k^T g(x_k) + \frac{M}{2} (\alpha'_k)^2 \|d_k\|^2 + \varepsilon_{k+1}, \end{aligned} \quad (4.4)$$

که در آن  $u_k = x_k + t \alpha'_k d_k, t \in (0, 1)$  و آخرین نابرابری از (۴,۱) پیروی می کند. از (۴,۴) پیروی می کند که

$$\begin{aligned} \frac{\alpha_k}{2} = \alpha'_k &> \left[ \frac{(g^\alpha(x_k, \varepsilon_k) - g(x_k))^T d_k - (1 - \sigma) g^\alpha(x_k, \varepsilon_k)^T d_k - \varepsilon_{k+1} / (\alpha'_k)^2}{\|d_k\|^2} \right] \frac{2}{M} \\ &\geq \left[ \frac{(1 - \sigma) \|g^\alpha(x_k, \varepsilon_k)\|^2 - \sqrt{2\varepsilon_k / \lambda} \|d_k\| - \varepsilon_k}{\|d_k\|^2} \right] \frac{2}{M} \\ &= \left[ \frac{(1 - \sigma) \|g^\alpha(x_k, \varepsilon_k)\|^2}{\|d_k\|^2} - o(\alpha_k) / \sqrt{\lambda} - o(\alpha_k^2) \right] \frac{2}{M} \end{aligned}$$

$$\geq \left[ \frac{(1 - \sigma)}{\left(1 + \frac{1}{c}\right)^2} - o(\alpha_k) / \sqrt{\lambda} - o(\alpha_k^2) \right] \frac{2}{M}, \quad (4.5)$$

که در آن، این تساوی از  $\varepsilon_k = o(\alpha_k^2 \|d_k\|^2)$  پیروی می کند، نابرابری دوم از (۲,۹)، (۳,۱۱) و  $\varepsilon_{k+1} \leq \varepsilon_k$  پیروی می کند، و آخرین نابرابری از (۳,۱۲) پیروی می کند.

$$\frac{1}{2} \geq \lim_{k \rightarrow \infty} \left( \frac{2(1 - \sigma)}{\left(1 + \frac{1}{c}\right)^2 M} \right) \frac{1}{\alpha_k} = +\infty.$$

این غیر ممکن است و بنابراین این نتیجه دنبال می شود.

ما اکنون جهت نشان دادن همگرایی کلی الگوریتم ۱ آماده هستیم.

**قضیه ۴,۱** (همگرایی کلی). فرض کنید که شرایط در گزاره ۴,۲ برقرار است. آنگاه،  $\lim_{k \rightarrow \infty} \|g(x_k)\| = 0$  و نقطه تجمع  $x_k$  یک راه حل بهینه از (۱,۱) است.

**اثبات.** ما نخست ثابت می کنیم که

$$\lim_{k \rightarrow \infty} \|g^\alpha(x_k, \varepsilon_k)\| = 0. \quad (4.6)$$

برای دیدن این، ما بوسیله روش تناقض ادامه می دهیم و فرض می کنیم که (۴,۶) شکست بخورد. بدون از دست دادن کلیت، ما شاید فرض کنیم که در اینجا ثابت های  $\varepsilon_0 > 0$  و  $k_0 > 0$  برآورده کننده

$$\|g^\alpha(x_k, \varepsilon_k)\| \geq \varepsilon_0, \quad \forall k > k_0. \quad (4.7)$$

وجود دارد.

با استفاده از (۳,۹)، (۳,۱۱)، (۴,۲) و (۴,۷)، ما بدست می آوریم

$$F^\alpha(x_{k+1}, \varepsilon_{k+1}) - J_k \leq \sigma \alpha_k g^\alpha(x_k, \varepsilon_k)^T d_k = -\sigma \alpha_k \|g^\alpha(x_k, \varepsilon_k)\|^2 \leq -\sigma m_0 \varepsilon_0, \quad \forall k > k_0.$$

از تعریف  $J_{k+1}$  دنبال می شود که

$$\begin{aligned} J_{k+1} &= \frac{\rho E_k J_k + F^\alpha(x_k + \alpha_k d_k, \varepsilon_{k+1})}{E_{k+1}} \\ &\leq \frac{\rho E_k J_k + J_k - \sigma m_0 \varepsilon_0}{E_{k+1}} \\ &= J_k - \frac{\sigma m_0 \varepsilon_0}{E_{k+1}}. \end{aligned} \quad (4.8)$$

از آنجا که  $F$  از زیر محدود شده است، ما می بینیم که  $F^\alpha(x, \varepsilon)$  از زیر محدود شده است. این همراه با  $F^\alpha(x, \varepsilon) \leq J_k$  برای تمام  $k$  ها نشان می دهد که  $J_k$  نیز از زیر محدود شده است. توسط (۴,۸)، ما داریم

$$\sum_{k=k_0}^{\infty} \frac{\sigma m_0 \varepsilon_0}{E_{k+1}} < \infty. \quad (4.9)$$

با تعریف  $E_{k+1}$  ما داریم  $E_{k+1} \leq k + 2$ . آن دنبال می شود که

$$\sum_{k=k_0}^{\infty} \frac{\sigma m_0 \varepsilon_0}{E_{k+1}} \geq \sum_{k=k_0}^{\infty} \frac{\sigma m_0 \varepsilon_0}{k+2} = +\infty.$$

این باعث تناقض می شود و بنابراین، (۴,۶) برقرار می شود.

ما اکنون ادعای دوم را نشان می دهیم. با استفاده از (۲,۹)، ما نخست می بینیم که

$$\|g^\alpha(x_k, \varepsilon_k) - g(x_k)\| \leq \sqrt{\frac{2\varepsilon_k}{\lambda}}.$$

این معادله با فرض (III) نشان می دهد که

$$\lim_{k \rightarrow \infty} \|g(x_k)\| = 0. \quad (4.10)$$

اجازه دهید  $x^*$  یک نقطه تجمع  $\{x_k\}$  باشد. آنگاه، در اینجا یک دنباله  $\{x_k\}_k$  برآورده کننده

$$\lim_{k \in K, k \rightarrow \infty} x_k = x^*. \quad (4.11)$$

وجود دارد.

بواسطه ویژگی های  $F(x)$ ، ما بدست می آوریم  $g(x_k) = (x_k - p(x_k))/\lambda$ . بنابراین بوسیله (۴,۱۰) و (۴,۱۱)،  $x^* = p(x^*)$  برقرار است. بنابراین  $x^*$  است.

قبل از اینکه ما این بخش را پایان دهیم، ما اظهار می کنیم که برخی روش های نوع گرادیان مزدوج در [۸,۱۰,۴۹] برای حل مسایل بهینه سازی غیر صاف ممکن بتازگی ارایه شدند. یکی از ویژگی های کلیدی روش مان که آن را از روش های پیشنهادی دیگر در [۸,۴۹] متمایز می کند، این است که روش ما نه تنها از اطلاعات گرادیان، بلکه از اطلاعات تابع نیز استفاده می کند.

## ۵. روش گرادیان مزدوج بیشتر اصلاح شده

مشابه به روش جدید (۳,۸) معرفی شده در بخش قبلی، ما می توانیم همان ایده و روش های مختلف گرادیان مزدوج اصلاح شده دیگر را برای حل مشکل (P) بکار گیریم. آنها به شرح زیر ذکر شده اند.

(۱) فرمول گرادیان مزدوج هستنزا-استیفیل (HS) برای یک مسئله بهینه سازی نامحدود صاف (نگاه کنید به [۴]):

$$d_{k+1} = \begin{cases} -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T y_k}{d_k^T y_k} d_k, & \text{if } k \geq 1, \\ -\nabla f_{k+1}, & \text{if } k = 0. \end{cases}$$

با استفاده از ایده مشابه در بخش قبل، می توان فرمول گرادیان مزدوج HS اصلاح شده زیر را برای حل مساله (P) پیشنهاد کرد:

$$d_{k+1} = \begin{cases} -g^\alpha(x_{k+1}, \varepsilon_{k+1}) + \frac{g^\alpha(x_{k+1}, \varepsilon_{k+1})^T y_k^* d_k - d_k^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) y_k^*}{\max\{2c \|d_k\| \|y_k^*\|, |d_k^T y_k^*|\}}, & \text{if } k \geq 1, \\ -g^\alpha(x_{k+1}, \varepsilon_{k+1}), & \text{if } k = 0. \end{cases}$$

فرمول گرادیان مزدوج لیو-استوری (LS) برای یک مسئله بهینه سازی نامحدود صاف (نگاه کنید به [۵۰]):

$$d_{k+1} = \begin{cases} -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T y_k}{-d_k^T \nabla f_k} d_k, & \text{if } k \geq 1, \\ -\nabla f_{k+1}, & \text{if } k = 0. \end{cases}$$

با استفاده از این ایده مشابه در بخش قبل، می توان فرمول گرادیان مزدوج LS اصلاح شده زیر را برای حل مساله (P) پیشنهاد کرد:

$$d_{k+1} = \begin{cases} -g^\alpha(x_{k+1}, \varepsilon_{k+1}) + \frac{g^\alpha(x_{k+1}, \varepsilon_{k+1})^T y_k^* d_k - d_k^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) y_k^*}{\max\{2c \|d_k\| \|y_k^*\|, |d_k^T g^\alpha(x_k, \varepsilon_k)|\}}, & \text{if } k \geq 1, \\ -g^\alpha(x_{k+1}, \varepsilon_{k+1}), & \text{if } k = 0. \end{cases}$$

(۳) فرمول گرادیان مزدوج دای-یوان (DY) برای مساله بهینه سازی نامحدود صاف ([۱] را ببینید):

$$d_{k+1} = \begin{cases} -\nabla f_{k+1} + \frac{\|\nabla f_{k+1}\|^2}{d_k^T y_k} d_k, & \text{if } k \geq 1, \\ -\nabla f_{k+1}, & \text{if } k = 0. \end{cases}$$

با استفاده از این ایده مشابه در بخش قبل، می توان فرمول گرادیان مزدوج DY اصلاح شده زیر را برای حل مساله (P) پیشنهاد کرد:

$$d_{k+1} = \begin{cases} -g^\alpha(x_{k+1}, \varepsilon_{k+1}) + \frac{\|g^\alpha(x_{k+1}, \varepsilon_{k+1})\|^2 d_k - d_k^T g^\alpha(x_{k+1}, \varepsilon_{k+1}) y_k^*}{\max\{2c \|d_k\| \|g^\alpha(x_{k+1}, \varepsilon_{k+1})\|, |d_k^T y_k^*|\}}, & \text{if } k \geq 1, \\ -g^\alpha(x_{k+1}, \varepsilon_{k+1}), & \text{if } k = 0. \end{cases}$$

مشابه الگوریتم ۱، ساخت الگوریتم های جدید و به دست آوردن همگرایی کلی این روش ها دشوار نیست. برای سادگی، ما جزئیات را در اینجا حذف کردیم.

## ۶. نتایج عددی

در این بخش، ما الگوریتم اصلاح شده PRP مان را (الگوریتم ۱) برای هر دو مسایل مقیاس کوچک و نیز مقیاس بزرگ امتحان می کنیم.

مسایل مقیاس کوچک. ما نخست الگوریتم مان را برای برخی از مسائل مقیاس کوچک امتحان می کنیم و آن را با روش بسته نزدیک در [۵۱] مقایسه می کنیم. همه مسائل غیرصاف جدول ۶،۱ را می توان در [۵۲] یافت. جدول ۶،۱ شامل ابعاد مساله و مقادیر تابع بهینه است.

در اینجا  $f_{ops}(x)$  مقدار تابع بهینه سازی است. الگوریتم با استفاده از نرم افزار Matlab 7.6 اجرا شد، و همه آزمایش ها بر روی یک کامپیوتر با پردازنده اینتل پنتیوم دو E7500 2.93 گیگاهرتز، حافظه ۲ گیگابایت SDRAM، و سیستم عامل ویندوز XP اجرا شدند.

پارامترها بعنوان  $\|g^\alpha(x, \varepsilon)\| \leq 10^{-10}$  برآورده شد، متوقف کردیم. به منظور نشان دادن عملکرد الگوریتم ارائه شده، ما همچنین نتایج مقاله [۵۱] (روش بسته نزدیک، PBL) و مقاله [۲۶] (مفهوم منطقه اعتماد، BT) را ذکر کردیم. نتایج عددی PBL و BT را می توان در [۵۱] یافت.



**Table 6.1**  
Problem descriptions for small-scaled testing problems.

Nr.	Problems	Dim	$f_{ops}(x)$
1	Rosenbrock	2	0
2	Crescent	2	0
3	CB2	2	1.9522245
4	CB3	2	2.0
5	DEM	2	-3
6	QL	2	7.20
7	LQ	2	-1.4142136
8	Mifflin 1	2	-1.0
9	Mifflin 2	2	-1.0
10	Rosen-Suzuki	4	-44
11	Shor	5	22.600162
12	Colville	5	-32.348679

**Table 6.2**  
Test results.

Nr.	Algorithm 1 NI/NF/ $f(x)$	PBL NI/NF/ $f(x)$	BT NI/NF/ $f(x)$	$f_{ops}(x)$
1	46/48/7.091824 $\times 10^{-7}$	42/45/0.381 $\times 10^{-6}$	79/88/0.130 $\times 10^{-11}$	0
2	11/13/6.735123 $\times 10^{-5}$	18/20/0.679 $\times 10^{-6}$	24/27/ $\times 10^{-6}$	0
3	12/14/1.952225	32/34/1.9522245	13/16/1.952225	1.9522245
4	2/6/2.000098	14/16/2.0	13/21/2.0	2.0
5	4/6/-2.999866	17/19/-3.0	9/13/-3.0	-3
6	10/12/7.200011	13/15/7.2000015	12/17/7.200009	7.20
7	2/3/-1.414214	11/12/-1.4142136	10/11/-1.414214	-1.4142136
8	4/6/-0.9919815	66/68/-0.99999941	49/74/-1.0	-1.0
9	20/23/-0.9999925	13/15/-1.0	6/13/-1.0	-1.0
10	28/58/-43.99986	43/45/-43.999999	22/32/-43.99998	-44
11	33/91/22.60023	27/29/22.600162	29/30/-22.60016	22.600162
12	17/23/-32.34329	62/64/-32.348679	45/45/-32.3486	-32.348679

**Table 6.3**  
Problem descriptions for large-scaled testing problems.

Nr.	Problems	$x_0$
1	Generalization of MAXQ	(1, 2, ..., n/2, -(n/2 + 1), ..., -n)
2	Generalization of MXHILB	(1, 1, ...)
3	Chained LQ	(-0.5, -0.5, ...)
4	Number of active faces	(1, 1, ...)
5	Nonsmooth generalization of Brown function 2	(1, 0, ...)
6	Chained Mifflin 2	(-1, -1, ...)
7	Chained Crescent I	(-1.5, 2, ...)
8	Chained Crescent II	(1, 0, ...)

ستون جدول ۶,۲ دارای معانی زیر است.

مساله: نام مساله آزمون. NF: تعداد ارزیابی های تابع.

NI: تعداد کل تکرار.  $f(x)$ : مقدار تابع در تکرار نهایی.

$f_{ops}(x)$ : ارزیابی تابع بهینه سازی.

از نتایج عددی در جدول ۶،۲، می توان مشاهده کرد که الگوریتم ۱ بصورت کلی در میان سه روش برای آزمایش مساله ذکر شده بهترین عملکرد را دارد و روش PBL نسبت به روش BT مقرون به صرفه است.

مسائل مقیاس بزرگ. همانطور که روش نوع گرادیان مزدوج به خصوص برای یک مساله مقیاس بزرگ مفید است، ما نیز برخی از آزمایش های عددی را برای مسائل محدب غیر صاف مقیاس بزرگ ارائه کردیم. به طور خاص، همانطور که روش موفق بسته حافظه محدود اخیر در [۲۷] ثابت کرده است که یکی از کارآمدترین روش ها برای حل یک مساله مقیاس بزرگ غیر صاف محدب است، ما الگوریتم مان را با آن مقایسه کردیم. مسائل بعدی جدول ۶،۳ را می توان در [۲۷] یافت، که در آن مسائل ۱-۳ تابع محدب و دیگران تابع غیر محدب هستند. تعداد متغیرهای مورد استفاده ۱،۰۰۰، ۵،۰۰۰، ۱۰،۰۰۰ و ۵۰،۰۰۰ بودند.

مقادیر پارامترها مشابه آنهایی هستند که در مسائل مقیاس کوچک بودند. آزمایشهای زیر در فرتن ۹۰ اجرا شدند. به منظور نشان دادن عملکرد الگوریتم ۱، ما آن را با روش (LMBM) مقاله [۲۷] مقایسه کردیم. قانون توقف و پارامترها در [۲۷] تنظیم شده اند.

LMBM [27]. روش بسته حافظه محدود جدید برای بهینه سازی غیر صاف مقیاس بزرگ. کدهای فرتن توسط Haarala، میتینن و Mäkelä، دخالت داده شدند، در <http://napsu.karmitsa.fi/lmbm/> در دسترس هستند.

برای این مسائل مقیاس بزرگ، تعداد تکرار و تعداد ارزیابی های تابع الگوریتم ۱ نسبت به آنهایی که از روش LMBM استفاده می کنند، مقرون به صرفه تر هستند. علاوه بر این، این تعداد، زمانی که بعد افزایش می یابد، بوضوح تغییر نمی کنند.

Table 6.4  
Test results.

Nr.	Dim	Algorithm 1 NI/NF/f (x)	LMBM NI/NF/f (x)
1	1000	$225/4710/6.93540513648899 \times 10^{-8}$	$21492/22259/6.71025884158921 \times 10^{-6}$
	5000	$250/5235/6.87979764297566 \times 10^{-8}$	$191470/196034/3.44987308783816 \times 10^{-5}$
	10000	$261/5466/6.65278880046151 \times 10^{-8}$	$512415/523351/5.83498629004021 \times 10^{-5}$
	50000	$286/5991/6.59944730089708 \times 10^{-8}$	$4999996/5000000/5.77622918225693 \times 10^2$
2	1000	$91/1482/8.27377560377777 \times 10^{-9}$	$441/861/6.16640522014495 \times 10^{-3}$
	5000	$111/1938/9.72057262375235 \times 10^{-9}$	$1258/2487/3.52143538240154 \times 10^{-2}$
	10000	$120/2127/5.85235371545923 \times 10^{-9}$	$7027/7810/5.11605598924595 \times 10^{-2}$
	50000	$141/2604/6.17794513724740 \times 10^{-9}$	$1036/1515/2.88716335417572$
3	1000	$37/114/7.26870782386942 \times 10^{-9}$	$300/1824/-1.41277614588146 \times 10^5$
	5000	$39/120/9.09316258227670 \times 10^{-9}$	$365/2198/-7.06961404239955 \times 10^5$
	10000	$40/123/9.09407298984668 \times 10^{-9}$	$376/2281/-1.41406858071499 \times 10^6$
	50000	$55/153/4.54740666059811 \times 10^{-8}$	$582/2998/-7.07092005296846 \times 10^6$
4	1000	$77/1026/6.80373977063602 \times 10^{-9}$	$523/569/1.37667655053518 \times 10^{-14}$
	5000	$90/1281/7.84048733972188 \times 10^{-9}$	$2585/2586/1.21306742421471 \times 10^{-10}$
	10000	$96/1401/9.93659283247496 \times 10^{-9}$	$5069/5073/5.38381117320365 \times 10^{-10}$
	50000	$110/1665/6.13431592528028 \times 10^{-9}$	$184/217/9.99980628711393 \times 10^6$
5	1000	$38/117/7.26870855073998 \times 10^{-9}$	$467/3873/4.05785228342877 \times 10^{-9}$
	5000	$40/123/9.09316349159170 \times 10^{-9}$	$453/4073/1.08041333809065 \times 10^{-8}$
	10000	$41/125/1.81881459796775 \times 10^{-8}$	$736/7453/2.52215161529255 \times 10^{-8}$
	50000	$55/153/9.09481332115074 \times 10^{-8}$	$1293/10995/7.26535476752977 \times 10^{-7}$
6	1000	$37/114/-2.49749999992731 \times 10^4$	$1254/7355/-7.06476909407459 \times 10^4$
	5000	$39/120/-1.24974999999091 \times 10^5$	$219/782/-3.53493693696815 \times 10^5$
	10000	$40/123/-2.49974999999091 \times 10^5$	$267/743/-7.0704203377240 \times 10^5$
	50000	$43/132/-1.24997499999986 \times 10^6$	$532/2220/-3.53546169719885 \times 10^6$
7	1000	$37/114/5.48971001990139 \times 10^{-9}$	$138/560/2.46289254738352 \times 10^{-4}$
	5000	$39/120/6.82939571561292 \times 10^{-9}$	$116/281/2.45751603804887 \times 10^2$
	10000	$40/123/6.82530298945494 \times 10^{-9}$	$188/267/2.00248216613019 \times 10^{-5}$
	50000	$56/157/8.52753601066070 \times 10^{-9}$	$391/725/4.75680539402390 \times 10^{-9}$
8	1000	$39/120/6.81848177919164 \times 10^{-9}$	$763/7522/1.39417095132655 \times 10^{-4}$
	5000	$41/126/8.52583070809487 \times 10^{-9}$	$943/8490/1.59176433435915 \times 10^{-3}$
	10000	$42/129/8.52617176860804 \times 10^{-9}$	$1364/13919/1.06009303953474 \times 10^{-2}$
	50000	$87/222/5.32902788563661 \times 10^{-9}$	$4657/61720/8.01042336673219 \times 10^{-4}$

مقدار تابع نهایی الگوریتم ارائه شده، بهتر از مقدار تابع LMBM است به جز برای مساله ۳. عملکرد نتایج عددی نشان می دهد که الگوریتم ۱ می تواند برای حل نه تنها مسائل بهینه سازی محدب صاف، بلکه همچنین برای مسائل بهینه سازی غیر صاف غیر محدب نیز استفاده شود. با بهره گیری از همه چیز، نتایج عددی مقدماتی نشان می دهند که روش ارائه شده، کارآمد است (جدول ۶،۴ را ببینید).

## 7. نتیجه گیری

روش CG سادگی دارد و حافظه بسیار کمی لازم است و روش PRP یکی از موثرترین روش های گرادیان مزدوج می باشد. با استفاده از تنظیم مورو-Yosida، یک تکنیک جستجو خط غیر یکنواخت [۴۸] و یک معادله قاطع جدید [۴۳] به دست آمده توسط نویسندگان پیشین، ما یک الگوریتم گرادیان مزدوج PRP اصلاح شده را برای حل مسائل بهینه سازی محدب غیر صاف ارائه کردیم. روش ما، ویژگی تنزل به اندازه کافی به طور خودکار را برآورده می کند، و جهت جستجو مربوطه متعلق به یک منطقه اعتماد است. یکی دیگر از ویژگی های جالب روش مان، این است که

آن نه تنها اطلاعات گرادیان بلکه اطلاعات تابع را نیز شامل می شود. نتایج عددی نشان می دهد که این روش، موثر است و نسبت به روش های موجود برای هر دو مسائل مقیاس کوچک و مقیاس بزرگ غیر صاف مقرون به صرفه تر است.

جالب خواهد بود که ببینید چگونه الگوریتم مان اجرا می شود اگر ما آن را برای حل بعضی از مسائل بهینه سازی که در منطقه پردازش تصویر ناشی می شوند، بکار ببریم. این نیز جالب خواهد بود که عملکرد روش مان را با روش ارائه شده در [۸،۱۰،۴۹] مقایسه کنیم. اینها موضوعات پژوهشی آینده ما خواهد بود و در یک مقاله در شرف بررسی قرار خواهد گرفت.

### سپاسگزاری

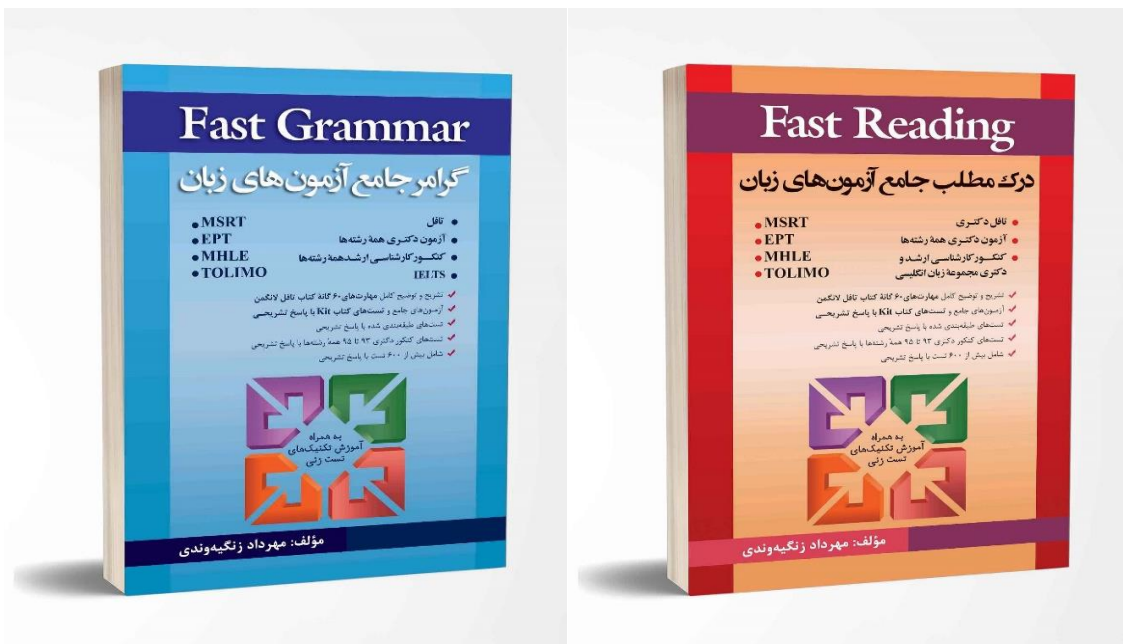
نویسندگان از داوران و سردبیر برای نظرات سازنده و پیشنهادات مفیدشان که در آماده سازی نهایی مقاله بکار برده شده اند، سپاسگزار می کنند.

Translated by DILMAJ Group

کاری از گروه ترجمه دیلماج

سجاد زنگیه وندی

وبسایت تخصصی FastZaban.com



وبسایت و فروشگاه تخصصی آزمون های زبان

[www.FastZaban.com](http://www.FastZaban.com)

دانلود منابع و جزوات آزمون های

MSRT, EPT, MHLE, TOEFL, IELTS

کارشناسی ارشد و دکتری همه رشته ها

فروش اینترنتی کتاب های تألیفی استاد

مهرداد زنگیه وندی

کتاب **Fast Grammar**

توضیح کامل ۶۰ مهارت لانگمن و گرامر آزمون های زبان و آموزش نحوه تست زنی در آزمون ها (در ۲۲۴ صفحه)  
به همراه سؤالات آزمون های سراسری و آزمون های MSRT, EPT, MHLE با پاسخ کاملاً تشریحی

کتاب **Fast Reading**

توضیح نکات و مهارت های درک مطلب آزمون های زبان و آموزش تکنیک های تست زنی در آزمون ها (در ۱۲۸ صفحه)  
به همراه سؤالات کنکور کارشناسی ارشد و دکتری همه رشته ها و آزمون های MSRT, EPT, MHLE با پاسخ تشریحی

خرید اینترنتی کتاب های فوق در:

[www.FastZaban.com](http://www.FastZaban.com)

کانال آموزش استاد مهرداد زنگیه وندی در تلگرام @FastZaban