بسم الله الرحمن الرحیم

دانشگاه صنعتی خواجه نصیر الدین طوسی

دانشکده علوم

گروه ریاضی

پایان‌نامه کارشناسی ارشد

گرایش تحقیق در عملیات

عنوان

پیرامون همگرایی روش ناحیه اعتماد گوس-نیوتنی غیردقیق برای حل مسائل حداقل مربعات غیرخطی با کران‌های ساده

نگارش

مرتضی غیاث‌الدین

استاد راهنما

دکتر محمد رضا پیغامی

استاد مشاور

دکتر مسعود فاطمی

شهریور ۱۳۹۲

دانشگاه صنعتی خواجه نصیر الدین طوسی

دانشکده علوم

نگارش: **مرتضی غیاث‌الدین**

استاد راهنما: **دکتر محمد رضا پیغامی** امضاء:

استاد ممتحن داخلی: امضاء:

استاد ممتحن خارجی: امضاء:

به پاس تعبیر عظیم و انسانی شان از کلمه ایثار

به پاس عاطفه سرشار و گرمای امید بخش وجودشان که در این سردترین روزگاران بهترین پشتیبان است

به پاس قلب های بزرگشان که فریادرس است و سرگردانی و ترس در پناهشان به شجاعت می گراید

و به پاس محبت های بی دریغشان که هرگز فروکش نمی کند

این مجموعه را به پدر و مادر عزیزم تقدیم می کنم.

چکیده

در این پایان‌نامه، یک روش ناحیه اعتماد غیردقیق برای حل مسایل کمترین مربعات غیرخطی مقید کران‌دار معرفی می‌شود. این روش مساله را به صورت یک مدل گوس–نیوتن در نظر گرفته و در هر تکرار، زیرمساله ناحیه اعتماد به صورت تقریبی توسط روش گرادیان مزدوج حل می‌شود. تحت شرایط کنترلی مناسب روی پارامترها، نشان داده می‌شود که روش پیشنهادی دارای ویژگی‌های همگرایی سراسری و مجانبی سریع است. در پایان، نتایج عددی، عملکرد روش را در عمل نیز تشریح می‌کنند.

کلمات کلیدی: بهینه‌سازی غیرخطی، روش ناحیه اعتماد، مساله کمترین مربعات غیرخطی، کران‌های ساده

سپاس بی‌کران پروردگار یکتا را که هستی‌مان بخشید و به طریق علم و دانش رهنمونمان شد و به هم‌نشینی رهروان علم و دانش مفتخرمان نمود و خوشه چینی از علم و معرفت را روزیمان ساخت.

من لم یشکر المخلوق لم یشکر الخالق

"هر کس از بندگان خداوند تشکر نکند از خداوند تشکر نکرده است"

اینک که به فضل خداوند این پایان نامه به سرانجام رسیده، وظیفه خود می‌دانم از زحمات بی دریغ اساتید و سروران گرامی که در اجرای این پایان‌نامه مرا یاری فرموده‌اند، به ویژه استاد ارجمند، جناب آقای دکتر محمدرضا پیغامی به خاطر راهنمایی‌ها و زحمات فراوانشان با سمت استاد راهنما و جناب آقای دکتر مسعود فاطمی با سمت استاد مشاور که در تمام این دوسال همواره مرا مورد لطف و محبت‌های بی دریغ و خالصانه خود قرار دادند، با تمامی وجود سپاسگزاری می‌کنم.

از آقایان دکتر اسماعیل خرم (داور خارجی) و دکتر محمد مسجدجامعی (داور داخلی) نیز به جهت خواندن رساله و پیشنهادات ارزنده‌شان کمال تشکر را دارم.

از تمامی دوستان خوبم به ویژه آقایان مصطفی دشت‌رزمی، نبی‌اله کرونی، بهنام رضایی، داوود عطایی، مظاهر حبیبی، سجاد فتحی، حمیدرضا ویسی، فرشید ارزانی و خانم زینب سعیدیان سپاسگزاری می‌کنم.

فهرست مطالب

فهرست تصاویر

فهرست جداول

فهرست الگوریتم‌ها

پیشگفتار

حل دستگاه معادلات و نامعادلات غیرخطی با کران ساده روی متغیرهای آن به دلیل کاربرد آن در زمینه‌هایی چون سیستم‌های تعادل شیمیایی، مسایل تعادل اقتصادی، مسایل مقدار مرزی، مسایل مکمل و مساله‌های بهینه‌سازی مقید از اهمیت زیادی برخوردار است.

وجود روش‌های بهینه‌سازی برای حل این دستگاه‌ها نیازمند یک مدل‌سازی خوب و مناسب از مساله اولیه می‌باشد، از این رو مدل‌سازی و الگوریتم‌های حل مسایل بهینه‌سازی غیرخطی جایگاه ویژه‌ای را در بین موضوعات و پژوهش‌های دانشمندان به خود اختصاص داده است. در این پایان‌نامه، ابتدا روش ناحیه اعتماد استاندارد[1] را مطرح و سپس نحوه مدل‌سازی مناسبی از سیستم‌های غیرخطی را شرح می‌دهیم. در نهایت اصلاحات انجام شده روی روش ناحیه اعتماد را برای حل سیستم‌های غیرخطی بیان می‌کنیم.

متناسب با اهداف پایان‌نامه، نحوه بیان مطالب در فصل‌های آن به صورت زیر تقسیم‌بندی شده‌اند. در فصل اول، مفاهیم اولیه و انواع مسایل بهینه‌سازی را بیان می‌کنیم. در فصل دوم، روش نیوتن را توضیح داده و به بیان ویژگی‌ها و مشکلات آن می‌پردازیم. در ادامه اصلاحات انجام شده روی آن‌را معرفی کرده و الگوریتم‌های مربوط به آن‌ها را ارایه می‌دهیم. در فصل سوم، روش ناحیه اعتماد استاندارد به همراه ویژگی‌های همگرایی آن‌را بیان و روش‌های حل زیرمساله را معرفی می‌کنیم. در فصل چهارم، ابتدا مساله کمترین مربعات غیرخطی را بیان و سپس روشی را که توسط پورچلی[2] [۲۵] برای حل این مساله ارایه شده است، مطرح می‌کنیم. این روش بر پایه روش ناحیه اعتماد عمل می‌کند. در فصل پنجم،

[1] Standard Trust Region
[2] Porcelli

نتایج عددی الگوریتم را در حالت‌های مختلف نشان می‌دهیم.

به طور کلی، بیشتر مطالب این پایان‌نامه از مراجع [۹]، [۱۹]، [۲۰]، [۲۲]، [۲۳]، [۲۵]، [۳۱]، [۳۲]، [۳۳]، [۳۴] برگرفته شده است.

فصل ۱

تعاریف و مقدمه‌ای بر بهینه‌سازی

۱.۱ مقدمه

به‌طور کلی بهینه‌سازی را علم یافتن ماکزیمم و مینیمم یک تابع با داشتن محدودیت‌هایی روی متغیرهای آن می‌دانیم. همواره در محیط پیرامون خود با مسایلی با این ویژگی روبرو می‌شویم، بنابراین سعی براین داریم که به نتیجه مطلوبی از حل مساله خود برسیم. یافتن جواب بهینه برای این‌گونه از مسایل که به مدل ریاضی تبدیل شده‌اند را در غالب مسایل بهینه‌سازی مطرح می‌کنیم. برای این منظور باید با شناخت مساله درپی مقادیر مناسب برای متغیرهای آن باشیم. با وجود محدودیت‌ها و شرایط مساله باید بتوان مساله را به بهترین شکل مدل‌بندی کرد تا به هدف خود که ماکزیمم یا مینیمم کردن است، برسیم. بنابراین مدل‌بندی مساله مهمترین مرحله در فرآیند بهینه‌سازی است. گاهی مدل‌بندی برای یک مساله به روش‌های مختلف انجام می‌شود که با توجه به الگوریتم‌های پیشنهادی، هرکدام می‌توانند متناظر با یک الگوریتم، مفید واقع شود.

از آن‌جایی که یک الگوریتم واحد برای تمام مسایل بهینه‌سازی وجود ندارد، پس از مدل‌بندی، انتخاب الگوریتم مناسب برای مساله مورد نظر، زمینه تحقیق و پژوهش دانشمندان شده است.

در ادامه این فصل تعاریف و مفاهیمی که در بهینه‌سازی مطرح و در این پایان‌نامه مورد استفاده قرار گرفته است، بیان می‌شوند.

در ابتدا نمادهای استفاده شده در این پایان‌نامه را بیان می‌کنیم. سپس مفهوم سرعت همگرایی توضیح

داده می‌شود. در ادامه فرم عمومی مساله بهینه‌سازی بیان و سپس مسایل بهینه‌سازی را دسته‌بندی می‌کنیم. در بخش بعد شرایط لازم و کافی بهینگی مطرح و در بخش آخر توابع محدب معرفی می‌گردند.

۲.۱ نمادگذاری

تعریف ۱.۲.۱. *فرض کنید تابع $f : \mathbb{R}^n \to \mathbb{R}$ داده شده باشد و $\Omega \subset \mathbb{R}^n$*

- *نقطه $x^* \in \Omega$ را یک نقطه مینیمم‌کننده موضعی f روی Ω گوییم هرگاه $\epsilon > \circ$ وجود داشته باشد به‌طوری‌که*

$$\forall x \in \Omega, |x - x^*| < \epsilon \qquad f(x^*) \leq f(x).$$

- *نقطه $x^* \in \Omega$ را یک نقطه مینیمم‌کننده موضعی اکید f روی Ω گوییم هرگاه $\epsilon > \circ$ وجود داشته باشد به‌طوری‌که*

$$\forall x^* \neq x \in \Omega, |x - x^*| < \epsilon \qquad f(x^*) < f(x).$$

- *نقطه $x^* \in \Omega$ را یک نقطه مینیمم‌کننده سراسری f روی Ω گوییم هرگاه*

$$\forall x \in \Omega \qquad f(x^*) \leq f(x).$$

تعریف ۲.۲.۱. *فرض کنید ماتریس $M \in S^n$، در این‌صورت تعاریف زیر را داریم:*

- *ماتریس M را معین مثبت[۱] گوییم، هرگاه:*

$$\forall d \in \mathbb{R}^n, \, d \neq \circ \; ; \; d^T M d > \circ.$$

[۱] Positive Definite

- *ماتریس M را نیمه معین مثبت[۲] گوییم، هرگاه:*

$$\forall d \in \mathbb{R}^n \ ; \ d^T M d \geq 0.$$

تعریف ۳.۲.۱. *تابع $f(\alpha)$ را از مرتبه $o(\alpha)$ (اوی کوچک) می‌گوییم هرگاه:*

$$\lim_{\alpha \to 0} \frac{f(\alpha)}{\alpha} = 0,$$

وتابع $f(\alpha) \geq 0$ را از مرتبه $O(\alpha)$ (اوی بزرگ) می‌گوییم هرگاه ثابت $c > 0$ وجود داشته باشد به‌طوری‌که:

$$\left| \frac{f(\alpha)}{\alpha} \right| \leq c.$$

در سرتاسر این پایان‌نامه، کلیه بردارها به‌صورت ستونی در نظر گرفته شده‌اند که درصورت لزوم توسط عملگر ترانهاده به سطری تبدیل می‌شوند.

$\|.\|$ نشان دهنده نرم l_2 برای بردار و ماتریس است. هم‌چنین $\|.\|_F$ و $\|.\|_M$ به ترتیب نشان دهنده نرم فروبینیوس[۳] و نرم وزن‌دار فروبینیوس[۴] هستند که به‌صورت زیر تعریف می‌شوند:

$$\|Q\|_F = \sqrt{Trace(QQ^T)},$$

$$\|Q\|_M = \|MQM\|_F,$$

که M یک ماتریس متقارن و معین مثبت است.

iـامین سطر ماتریس A را با $(A)_i$ وiـامین مولفه بردار x را با $(x)_i$ نشان می‌دهیم. A^+، ماتریس شبه وارون موره‌ـپنرز[۵] ماتریس A را نشان می‌دهد. به ازای ثابت $\rho > 0$، گوی به مرکز y و به شعاع

[۲] Positive Semidefinite

[۳] Frobenius norm

[۴] Weighted Frobenius norm

[۵] Moore-Penrose

ρ عبارت است از:

$$B_\rho(y) = \{x : \|x - y\| < \rho\}.$$

تعریف ۴.۲.۱. *فرض کنید تابع $f(x)$ دوبار مشتق پذیر باشد، در این‌صورت بردار گرادیان تابع f و عملگر ∇ را به‌صورت زیر تعریف می‌کنیم*

$$\nabla = \left(\frac{\partial}{\partial x_۱}, ..., \frac{\partial}{\partial x_n}\right)^T \quad , \quad \nabla f(x) = \left(\frac{\partial f}{\partial x_۱}, ..., \frac{\partial f}{\partial x_n}\right)^T,$$

هم‌چنین ماتریس هسی این تابع به‌صورت زیر تعریف می‌گردد:

$$\nabla^۲ f = \nabla \left[\frac{\partial f}{\partial x_۱}, ..., \frac{\partial f}{\partial x_n}\right]^T = \begin{pmatrix} \frac{\partial^۲ f}{\partial {x_۱}^۲} & \cdots & \frac{\partial^۲ f}{\partial x_۱ \partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial^۲ f}{\partial x_n \partial x_۱} & \cdots & \frac{\partial^۲ f}{\partial {x_n}^۲} \end{pmatrix}.$$

تعریف ۵.۲.۱. *تابع $[t]_+$ به صورت زیر تعریف می‌شود:*

$$[t]_+ = max\{t, ۰\}.$$

واضح است که $[t]_+$ به‌طور پیوسته مشتق پذیر است:

$$\frac{d}{dt}\left(\frac{۱}{۲}[t]_+^۲\right) = [t]_+.$$

۳.۱ همگرایی و مرتبه همگرایی

بیشتر روش‌های بهینه‌سازی بر پایه الگوریتم‌های تکراری هستند، که با تولید دنباله $\{x_k\}$ به سمت جواب حرکت می‌کنند. حال بررسی این‌که آیا یک الگوریتم به جواب مساله همگرا می‌شود یا خیر، موضوع مهمی است که برای سنجش کارایی الگوریتم مورد بحث قرار می‌گیرد. بنابراین همگرایی تنها یک مشخصه از الگوریتم نیست، بلکه خاصیتی است که به روش حل مساله، از دیدگاه نظری اعتبار می‌بخشد.

برای درک بهتر همگرایی، دو مفهوم زیر را بیان می‌کنیم:

- همگرایی سراسری: با شروع از هرنقطه آغازین دلخواه، دنباله نقاط تولید شده $\{x_k\}$ توسط روش حل مساله به جواب همگرا می‌شود.
- همگرایی موضعی یا مجانبی: با شروع از نقطه آغازین به اندازه کافی نزدیک به جواب، دنباله نقاط تولید شده $\{x_k\}$ توسط روش حل مساله به جواب همگرا می‌شود.

۱.۳.۱ مرتبه همگرایی

برای انتخاب یک روش مناسب از بین روش‌هایی که دارای همگرایی هستند، می‌توان از شاخص مرتبه همگرایی که مرتبط با سرعت همگرایی است، کمک گرفت که در ادامه به آن می‌پردازیم.

تعریف ۱.۳.۱. *فرض کنید $\{x_k\}$ دنباله ای همگرا به x^* باشد، مرتبه همگرایی دنباله فوق را p^* گوئیم هرگاه:*

$$p^* = \sup\{p : \beta = \limsup \frac{\|x_{k+1} - x^*\|_Q}{\|x_k - x^*\|_Q^p} < \infty\}.$$

توجه کنید که هرچه p^* بزرگتر باشد سرعت همگرایی نیز بیشتر خواهد بود، بنابراین داریم:

۱. اگر $p^* = 1$ و $0 < \beta < 1$، آنگاه همگرایی Q ـ خطی[۶] نامیده می‌شود.

۲. اگر $p^* = 1$ و $\beta = 0$، آنگاه همگرایی Q ـ زبر خطی [۷] نامیده می‌شود.

۳. اگر $p^* = 1$ و $\beta = 1$، آنگاه همگرایی Q ـ زیر خطی[۸] نامیده می‌شود.

۴. اگر $p^* = 2$ آنگاه همگرایی Qـمرتبه دوم [۹] نامیده می‌شود.

۵. اگر $p^* = 3$ آنگاه همگرایی Qـمکعبی [۱۰] نامیده می‌شود.

[۶] linear
[۷] Superlinear
[۸] Sublinear
[۹] Quadratic
[۱۰] Cubic

بنابراین می‌توان گفت که در مقایسه چند روش، روشی مناسب‌تر است که دارای مرتبه همگرایی بزرگتری است، البته مشروط بر این‌که روش‌ها از نظر پیچیدگی محاسباتی (حجم محاسبات لازم برای تولید تکرار جدید) در هر تکرار تفاوت معناداری نداشته باشند.

۴.۱ فرم عمومی مساله بهینه‌سازی

فرم عمومی یک مساله بهینه‌سازی به‌صورت زیر است:

$$\begin{aligned}
\min \quad & f(x) \\
s.t. \quad & c_i(x) = \circ \quad ; i \in \mathcal{E} \\
& c_j(x) \geq \circ \quad ; j \in \mathcal{I} \\
& x \in C
\end{aligned} \tag{۱.۱}$$

که در آن $\mathcal{E} = \{۱, ..., p\}$ و $\mathcal{I} = \{۱, ..., m\}$ به ترتیب مجموعه اندیس‌های قیود مساوی و نامساوی هستند، C یک مجموعه محدب دلخواه است و توابع c_i و c_j روی C تعریف می‌گردند.
مساله (۱.۱) را یک مساله بهینه‌سازی غیرخطی می‌نامند هرگاه حداقل یکی از توابع $f(x)$ ، c_i یا c_j غیرخطی باشند.

تعریف ۱.۴.۱. *مجموعه جواب‌های شدنی مساله (۱.۱) را به‌صورت زیر تعریف می‌کنیم:*

$$F = \{x | c_i(x) = \circ, c_j(x) \geq \circ \quad \forall i, \forall j\}.$$

در این‌صورت:

- اگر $F = \varnothing$ ، آنگاه مساله را نشدنی گوییم.
- اگر تابع $f(x)$ روی F کران‌دار نباشد، آنگاه مساله را نامحدود گوییم.
- اگر $minf(x)$ در $x^* \in F$ به‌دست آید، x^* را جواب بهینه مساله گوییم.