

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده علوم ریاضی

رساله دکتری تحقیق در عملیات

روش‌هایی برای حل مسایل مکان‌یابی توافقی

نگارنده: آریا سلیمانی کورنده

استاد راهنما

دکتر جعفر فتحعلی

استاد مشاور

دکتر احمد نزاکتی

شهریور ۱۴۰۰

تقدیم به

پدر و مادر کراتقدم...

سپاس‌گزاری...

اینک که به فضل خداوند این رساله به سرانجام رسیده است، شایسته است با احترام از زحمات بی‌دریغ استاد ارجمندم، جناب آقای دکتر فتحعلی که همواره راهنما و حامی بنده بوده‌اند، تشکر نمایم. همچنین بر خود لازم می‌دانم که از راهنمایی‌های جناب آقای دکتر نزاکتی سپاس‌گزاری نمایم. همچنین، قدردان زحمات اساتید گرامی جناب آقای دکتر غزنوی و جناب آقای دکتر نوری اسکندری از دانشگاه صنعتی شاهرود و جناب آقای دکتر بابایی کفاکی از دانشگاه سمنان هستم که زحمت مطالعه و داوری این رساله را به عهده گرفته‌اند. سپاس‌گزار خانواده مهربانم به ویژه پدر و مادر عزیزم هستم که همواره یار و مشوق من هستند و هر چه دارم از آن‌هاست. همچنین از سرکار خانم سارا طاهری فرد که در طول مدت انجام این رساله همواره حامی و پشتیبان بنده بوده‌اند کمال تشکر را دارم.

آریا سلیمانی کورنده

شهریور ۱۴۰۰

تعهد نامه

اینجانب آریا سلیمانی کورنده دانشجوی دکتری رشته ریاضی کاربردی علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان **روش‌هایی برای حل مسایل مکان‌یابی توافقی**، تحت راهنمایی **جعفر فتحعلی** متعهد می‌شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام “دانشگاه صنعتی شاهرود” یا “Shahrood University of Technology” به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

آریا سلیمانی کورنده

شهریور ۱۴۰۰

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

نظریه مکان‌یابی یکی از مباحث مهم در بهینه‌سازی و تحقیق در عملیات است. در مسایل مکان‌یابی، هدف پیدا کردن مکان یک یا چند سرویس‌دهنده به گونه‌ای است که معیارهایی مانند هزینه حمل‌ونقل، مسافت طی‌شده توسط مشتریان، زمان کل سرویس‌دهی و هزینه ناشی از آن بهینه شود. در این رساله، به حل مسایل مکان‌یابی توافقی می‌پردازیم که در آن، مکان تعدادی مشتری در صفحه داده شده است و حالت ایده‌آل این است که مکان سرویس‌دهنده را چنان تعیین کنیم که فاصله سرویس‌دهنده تا مشتری i - ام برابر R_i باشد، اما چون چنین جوابی همواره موجود نیست، به دنبال کمینه کردن مجموع خطای حاصل از فاصله سرویس‌دهنده تا نقطه ایده‌آل هستیم. بدین منظور، دو مدل تابع زیان (خطا) متقارن و نامتقارن در نظر می‌گیریم و از تابع فاصله نرم ℓ_p در تعریف توابع زیان استفاده می‌کنیم. به منظور کمینه‌سازی تابع زیان متقارن از الگوریتم‌های شبه‌وایزفیلد، گاوس-نیوتن و الگوریتم فراابتکاری رقابت استعماری استفاده می‌کنیم و نتایج عددی حاصل از حل الگوریتم‌های ارائه‌شده را با هم مقایسه می‌کنیم. سپس در ادامه خصوصیات تابع زیان لاینکس را بیان کرده و مساله مکان‌یابی توافقی با این تابع زیان را بررسی می‌کنیم. به کمک روش‌های اصلاح‌شده جدید BFGS، الگوریتم‌های رقابت استعماری و شبه وایزفیلد به کمینه‌سازی این تابع زیان می‌پردازیم و نتایج حاصل از پیاده‌سازی این الگوریتم‌ها را ارائه داده و این نتایج را با یکدیگر مقایسه می‌کنیم. هم‌چنین در قسمت نهایی این رساله در پی حل مسایل مقیاس بزرگ با استفاده از روش‌های بهینه‌سازی تصادفی هستیم. روش‌های بهینه‌سازی تصادفی با کاهش هزینه محاسبات در مسایل مقیاس بزرگ عملکرد بسیار مطلوبی از خود نشان می‌دهند و در زمان کمتری به مقدار تابع هدف کمتر دست می‌یابند که همین موضوع موجب محبوبیت این روش‌ها در میان محققان حوزه بهینه‌سازی شده است. ما نیز با به کارگیری روش‌های تصادفی به حل مساله مکان‌یابی توافقی تحت تابع زیان متقارن پرداخته‌ایم و نتایج بسیار مطلوبی به دست آورده‌ایم که این نتایج را در قسمت پایانی رساله با استفاده از نمودارها و جداول مختلفی ارائه و شرح داده‌ایم.

کلمات کلیدی: مکان‌یابی توافقی، تابع زیان، الگوریتم فراابتکاری رقابت استعماری، روش شبه وایزفیلد، روش‌های BFGS، روش‌های بهینه‌سازی تصادفی.

لیست مقالات مستخرج از پایان نامه

۱. آریا سلیمانی کورنده، جعفر فتحعلی، (۱۳۹۴)، روش‌های شبه وایزفیلد برای مساله مکان‌یابی توافقی تحت نرم ℓ_p ، اولین کنفرانس ملی بهینه‌سازی و تصمیم‌گیری، دانشگاه مازندران.

۲. آریا سلیمانی کورنده، جعفر فتحعلی، (۱۳۹۷)، حل مساله مکان‌یابی تک‌وسیله‌ای توافقی تحت تابع زیان نامتقارن لاینکس، دومین سمینار ملی کنترل و بهینه‌سازی، دانشگاه صنعتی شاهرود.

۳. آریا سلیمانی کورنده، جعفر فتحعلی، مرتضی نظری، (۱۳۹۷)، مسایل مکان‌یابی تک‌وسیله‌ای آرمانی تحت نرم ℓ_p ، پژوهش‌های نوین در تصمیم‌گیری، دوره ۳، شماره ۴.

4. Soleimani A., Fathali J., Nezakati A., Nazari M. (2019). "Single Facility Goal Location Problems with Symmetric and Asymmetric Penalty Functions", Control and Optimization in Applied Mathematics (COAM), 4, pp 1–14.
5. Soleimani A., Taherifard S., Fathali J., (2021). "Solving single facility goal Weber location problem using stochastic optimization methods", Iranian Journal of Operations Research, submitted.

فهرست مطالب

س	فهرست تصاویر
ف	فهرست جداول
ق	پیشگفتار
۱	۱ مقدمه‌ای بر مفاهیم بهینه‌سازی
۱	۱.۱ مفاهیم مقدماتی
۴	۲.۱ ساختار روش‌های عددی برای حل مسایل بهینه‌سازی
۵	۱.۰.۲.۱ ساختار روش‌های مبتنی بر جستجوی خطی
۷	۲.۰.۲.۱ تعیین جهت جستجوی خطی
۸	۳.۰.۲.۱ روش تندترین کاهش
۸	۴.۰.۲.۱ روش نیوتن
۹	۵.۰.۲.۱ روش‌های شبه نیوتن
۱۱	۶.۰.۲.۱ روش BFGS حافظه محدود
۱۲	۷.۰.۲.۱ روش بی‌حافظه BFGS
۱۴	۸.۰.۲.۱ روش گاوس-نیوتن
۱۷	۲ مقدمه‌ای بر مفاهیم مکان‌یابی
۱۸	۱.۲ زمینه‌های عمده تحقیق در نظریه مکان‌یابی
۲۱	۲.۲ مولفه‌های اساسی مسایل مکان‌یابی
۲۱	۱.۲.۲ فضای مسایل مکان‌یابی
۲۳	۲.۲.۲ سرویس‌دهندگان
۲۴	۳.۲.۲ اندازه‌گیری فاصله
۲۵	۴.۲.۲ وزن
۲۵	۵.۲.۲ مشتریان
۲۵	۶.۲.۲ اهداف تصمیم‌گیرنده

۳.۲	تابع هدف‌های پرکاربرد مکان‌یابی	۲۶
۴.۲	برخی مسایل مهم مکان‌یابی	۲۶
۱.۴.۲	مساله مکان‌یابی چند وسیله‌ای پیوسته	۲۷
۲.۴.۲	مساله p -میان و مدل ریاضی آن	۲۸
۳.۴.۲	مساله مکان‌یابی p - مرکز	۲۸
۳	مسایل مکان‌یابی تک‌وسيله‌ای توافقی تحت تابع زیان متقارن	۳۱
۱.۳	مقدمه	۳۱
۲.۳	مبانی نظری و پیشینه تحقیق	۳۲
۳.۳	مساله مکان‌یابی تک وسیله‌ای توافقی با نرم ℓ_p	۳۴
۱.۳.۳	بیان مساله	۳۴
۲.۳.۳	مدل‌سازی مساله	۳۴
۴.۳	الگوریتم‌های حل مساله مکان‌یابی تک وسیله‌ای توافقی	۳۷
۱.۴.۳	الگوریتم شبه‌وایزفیلد	۳۷
۲.۴.۳	الگوریتم گاوس- نیوتن	۴۰
۳.۴.۳	الگوریتم فراابتکاری رقابت استعماری [۵۸]	۴۰
۵.۳	نتایج عددی	۴۲
۴	مسایل مکان‌یابی وبر توافقی تحت تابع زیان نامتقارن	۴۷
۱.۴	مساله مکان‌یابی وبر توافقی تحت تابع زیان نامتقارن لاینکس	۴۹
۱.۱.۴	روش شبه- وایزفیلد	۵۰
۲.۱.۴	روش BFGS و نسخه‌های اصلاح‌شده آن	۵۱
۲.۴	نتایج عددی	۵۲
۵	حل مساله مکان‌یابی توافقی در مقیاس بزرگ با رویکرد تصادفی	۵۷
۱.۵	مقدمه	۵۷
۲.۵	روش‌های بهینه‌سازی تصادفی	۵۹
۱.۲.۵	روش‌های مرتبه اول	۵۹
۱.۱.۲.۵	روش گرادیان کاهشی تصادفی (SGD)	۵۹
۲.۱.۲.۵	روش‌های کاهش واریانس	۶۰
۲.۲.۵	روش‌های مرتبه دوم	۶۱
۱.۲.۲.۵	روش‌های نیوتن بدون استفاده از ماتریس هسی	۶۲
۲.۲.۲.۵	روش‌های شبه نیوتن	۶۳
۳.۲.۵	روش‌های نیوتن منظم‌سازی شده مکعبی	۶۳
۴.۲.۵	برخی از روش‌های متداول دیگر	۶۴

۶۵	۱.۴.۲.۵	روش‌های گرادیان با گشتاور	۶۵
۶۵	۲.۴.۲.۵	روش‌های گرادیان تسریع شده	۶۵
۶۶	۳.۴.۲.۵	روش‌های مختصات کاهشی	۶۶
۶۶	۳.۵	نسخه تصادفی الگوریتم LBFGS	۶۶
۶۷	۱.۳.۵	الگوریتم SLBFGS	۶۷
۷۱	۴.۵	نتایج عددی	۷۱
۸۱		نتیجه‌گیری و پیشنهادات آینده	۸۱
۸۳		مراجع	۸۳
۹۷		واژه‌نامه فارسی به انگلیسی	۹۷
۱۰۱		واژه‌نامه انگلیسی به فارسی	۱۰۱

فهرست تصاویر

۶	۱.۱	طول گام‌های صادق در شرایط ولف [۱۰۵]
۷	۲.۱	طول گام‌های صادق در شرایط گلدشتاین [۱۰۵]
۸	۳.۱	روش تندترین کاهش [۱۰۵]
۳۵	۱.۳	نمای گرافیکی مساله مکان‌یابی توافقی
	۲.۳	دو نمای گوناگون از تابع هدف مساله (۸.۳) با داده‌های جدول ۳.۳ و شعاع
۴۴		توافقی I
	۳.۳	نمودار کاهشی گرادیان تابع هدف الگوریتم GN برای مثال جدول ۳.۳ با
۴۵		شعاع توافقی II
	۴.۳	نمودار کاهشی گرادیان تابع هدف الگوریتم ICA برای مثال جدول ۵.۳ با
۴۵		شعاع توافقی II
۴۸	۱.۴	نمودار تابع زیان لاینکس نسبت به مقادیر مختلف a در حالت $b = 1$
	۲.۴	نمودار دولان – دوره بر حسب زمان اجرا برای مقایسه کارایی الگوریتم‌های
۵۶		مورد آزمون
	۱.۵	مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های
۷۲		SVRG، SGD و SLBFGS بر حسب تکرارها برای مثال با ۳۰ عضو
	۲.۵	مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده
۷۲		از الگوریتم‌های SVRG و SLBFGS بر حسب تکرارها برای مثال با ۳۰ عضو
	۳.۵	مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های
۷۳		SVRG، SGD و SLBFGS بر حسب تکرارها برای مثال با ۱۰۰ عضو
	۴.۵	مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده
۷۳		از الگوریتم‌های SVRG و SLBFGS بر حسب تکرارها برای مثال با ۱۰۰ عضو
	۵.۵	مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های
۷۴		SVRG، SGD و SLBFGS بر حسب تکرارها برای مثال با ۱۰۰۰ عضو

- ۶.۵ مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده
 ۷۵ از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو
- ۷.۵ مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های
 ۷۷ SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۳۰ عضو و $m = 10$
- ۸.۵ مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده
 ۷۷ از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۳۰ عضو
 و $m = 10$
- ۹.۵ مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های
 ۷۸ SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰ عضو و $m = 10$
- ۱۰.۵ مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده
 ۷۸ از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰ عضو
 و $m = 10$
- ۱۱.۵ مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های
 ۷۹ SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو و $m = 32$
- ۱۲.۵ مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده
 ۷۹ از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو
 و $m = 32$

فهرست جداول

۴۲	مولفه‌های تنظیم‌شده الگوریتم (ICA)	۱.۳
۴۳	مختصات نقاط تقاضا $A_i = (a_i, b_i)$ برای حالت $p = ۲$ و $w_i = R_i = ۱$	۲.۳
۴۴	داده‌های یک مساله با ۴ نقطه برای حالت $p = ۲$	۳.۳
۴۴	داده‌های یک مساله با ۳۰ نقطه برای $R_i = ۱$	۴.۳
۴۴	نتایج به‌دست‌آمده از داده‌های جدول ۳.۳ برای $p = ۲$	۵.۳
۴۵	نتایج به‌دست‌آمده از داده‌های جدول ۴.۳ نسبت به مقادیر مختلف p	۶.۳
۴۶	نتایج حاصله از الگوریتم‌های GN و ICA نسبت به مقدار مختلف n و p	۷.۳
۵۲	نسخه‌های اصلاح شده روش BFGS و نویسندگان آن‌ها	۱.۴
۵۳	مختصات، وزن‌ها و شعاع توافقی نقاط موجود	۲.۴
۵۳	زمان اجرای روش‌ها	۳.۴
۵۳	نتایج حاصله از BFGS برای مقادیر مختلف p و شعاع توافقی دوم	۴.۴
۵۴	نتایج حاصله از مثال‌های مقیاس بزرگ برای مقادیر مختلف p	۵.۴
۵۴	زمان اجرای چند تکرار برای مثال‌های در مقیاس بزرگ برای مقادیر مختلف p	۶.۴
۵۹	خلاصه‌ای از روش‌های بهینه‌سازی تصادفی	۱.۵
۷۵	مقادیر بهینه، جواب بهینه و سرعت اجرای الگوریتم‌های SGD، SVRG و SLBFGS برای مقادیر مختلف p	۲.۵
۷۵	مقادیر بهینه، جواب بهینه، سرعت اجرا و واریانس برآوردگر گرادیان الگوریتم‌های SVRG، SLBFGS و SGD برای مقادیر مختلف m	۳.۵

پیشگفتار

یکی از مسایل مهمی که همواره مورد توجه و چالش پژوهشگران و محققان دنیا بوده است، مساله مکان‌یابی نام دارد. مساله مکان‌یابی به انتخاب موقعیت مناسب برای استقرار تسهیلات می‌پردازد به گونه‌ای که بتوان با نیازهای خاصی هماهنگی مطلوبی ایجاد نمود. نیازهای مربوط به استقرار و انتخاب موقعیت محل تسهیلات با توجه به معیارهای موردنظر، مکان‌یابی نامیده می‌شود.

مسایل مربوط به نظریه مکان‌یابی با توجه به مؤلفه‌های مهمی از قبیل حضور مستمر در زندگی و برنامه‌ریزی انسانی، ماهیت استراتژیک و به‌کارگیری منابع عظیمی از سرمایه، تأثیرات اقتصادی بلندمدت و پیچیدگی محاسباتی، در چندین دهه گذشته توجه ویژه‌ای را به خود اختصاص داده‌اند.

مکان‌یابی تسهیلات و تأسیسات از کاربردی‌ترین زمینه‌های تحقیقاتی است که گستره آن بسیاری از شاخه‌های علوم را در بر می‌گیرد. از جمله این شاخه‌ها می‌توان به تحقیق در عملیات، علوم مدیریت، مهندسی صنایع، جغرافیا و دیگر شاخه‌های مرتبط اشاره کرد. توجه به مسایل مکان‌یابی، سبب کاهش هزینه‌ها و موفقیت واحدهای صنعتی می‌شود؛ به‌طوری‌که مکان‌یابی یک فعالیت، اعم از یک بنگاه خرده‌فروشی، کارخانه، مرکز خدماتی و غیره، از مهم‌ترین پرسش‌های پیش روی یک فعالیت اقتصادی است تا آنجا که این مساله می‌تواند تأثیرات بسزایی در موفقیت یا شکست آن داشته باشد.

در طول دهه‌های اخیر، توجه و علاقه‌مندی به مسایل محیطی و دنیای واقعی کاربردی افزایش پیدا کرده است. این مسایل عمدتاً در نتیجه پیشرفت‌های صنعتی و فناوری ایجاد شده‌اند و غالباً تأثیرات منفی در محیط و مشکلات زیست‌محیطی را به دنبال می‌آورند.

مسایلی مانند استقرار یک سرویس‌دهنده که در عین تأمین خدمات و تقاضای مشتریان، اثرات نامطلوبی مثل سروصدا، گردوخاک، گازهای سمی، بوی بد، شلوغی، تشعشعات هسته‌ای و غیره را بر مشتریان تحمیل می‌کند، از این دسته هستند. تصمیم‌گیری در مورد احداث چنین سرویس‌دهندگانی در دنیای واقعی و زندگی، کار پیچیده و دشواری بوده و تعداد آن‌ها نیز قابل توجه است. مکان‌یابی چنین سرویس‌دهندگانی که هم دارای تأثیرات مثبت و هم تأثیرات منفی بر روی مشتریان هستند، مکان‌یابی توافقی نام‌گذاری شده است.

بنابراین با توجه به اهمیت موضوع و وجود بسیاری از مسایل در دنیای واقعی که با استفاده

از مفهوم مکان‌یابی توافقی به‌صورت ریاضی مدل می‌شوند، در این رساله به بررسی مسایل مکان‌یابی تک‌وسیله‌ای توافقی پرداخته می‌شود.

در فصل اول به توضیح مفاهیم و تعاریف لازم علم بهینه‌سازی می‌پردازیم و قضایا و روش‌های مورد استفاده را عنوان می‌کنیم. در فصل دوم مقدمه‌ای از مفاهیم اولیه علم مکان‌یابی را بیان کرده و زمینه‌های عمده تحقیقاتی در نظریه مکان‌یابی را معرفی می‌کنیم. در فصل سوم به حل مساله مکان‌یابی تک‌وسیله‌ای توافقی تحت تابع زیان متقارن مربعات خطا و قدرمطلق می‌پردازیم و با استفاده از چندین روش بهینه‌سازی، تابع زیان مربوطه را کمینه کرده و نتایج عددی مرتبط را تحت نمودارها و جداول متعدد ارائه می‌دهیم. فصل چهارم به حل مساله مکان‌یابی توافقی تحت تابع زیان نامتقارن لاینکس به کمک تابع فاصله نرم ℓ_p می‌پردازد و با استفاده از الگوریتم‌های جدید اصلاح‌شده BFGS، شبه وایزفیلد و الگوریتم فراابتکاری رقابت استعماری مساله مذکور را حل می‌کند. در این فصل نیز نتایج عددی مطلوبی حاصل شده است که این نتایج را در بخش پایانی ارائه می‌دهیم. به‌علاوه، فصل آخر دارای رویکردی تصادفی برای حل مسایل مکان‌یابی توافقی در مقیاس بزرگ است که با استفاده از روش‌های بهینه‌سازی تصادفی نتایج بسیار مطلوبی به‌دست می‌آید. روش‌های بهینه‌سازی تصادفی با کاهش هزینه محاسباتی مسایل مقیاس بزرگ موجب افزایش سرعت اجرای الگوریتم‌ها می‌شوند، به‌طوری‌که یک الگوریتم تصادفی در زمان کمتری به مقدار تابع زیان کمتر دست می‌یابد که همین موضوع توجه بسیاری از محققان را در سال‌های اخیر به خود جلب کرده است. در این رساله به معرفی چندین تکنیک بهینه‌سازی تصادفی می‌پردازیم سپس با استفاده از نسخه تصادفی روش حافظه محدود BFGS، تابع زیان مساله مورد نظر را کمینه می‌کنیم. نتایج عددی حاصله را ارائه می‌دهیم و این نتایج را با نتایج چند الگوریتم بهینه‌سازی تصادفی دیگر مقایسه کرده و در قالب نمودارها و جداول گوناگون ارائه می‌دهیم.

فصل ۱

مقدمه‌ای بر مفاهیم بهینه‌سازی

در این فصل، مفاهیم و روش‌های بهینه‌سازی کلاسیک را بیان می‌کنیم که این مطالب از مرجع [۱۰۵] برگرفته شده‌اند.

۱.۱ مفاهیم مقدماتی

تعریف ۱.۱.۱. تابع حقیقی $f : C \rightarrow [0, +\infty]$ با ضابطه $f(x) = \|x\|$ را نرم نامند، هر گاه در شرایط زیر صدق کند:

$$\|x\| = 0 \Leftrightarrow x = 0, \|x\| \geq 0 \quad \bullet$$

$$\alpha \in \mathbb{R} \quad \text{که} \quad \|\alpha x\| = |\alpha| \|x\| \quad \bullet$$

$$\|x + y\| \leq \|x\| + \|y\| \quad \bullet$$

برای $C = \mathbb{R}^n$ ، متداول‌ترین نرم‌ها، نرم برداری ℓ_p است، که برای $x \in \mathbb{R}^n$ به صورت زیر تعریف می‌شود:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad p \in \mathbb{R}^+$$

رایج‌ترین نرم‌های برداری ℓ_p ، برای $x \in \mathbb{R}^n$ نرم‌های ℓ_1 ، ℓ_2 و ℓ_∞ هستند که به صورت زیر نمایش داده می‌شوند:

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \sqrt{x^T x}, \quad \|x\|_\infty = \max_{i \in \{1, \dots, n\}} |x_i|$$

نرم ℓ_2 را نرم اقلیدسی^۱ می‌نامند.

تعریف ۲.۱.۱. تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ را در نقطه x مشتق‌پذیر گویند، هرگاه يك تابع $g: \mathbb{R}^n \rightarrow \mathbb{R}$ موجود باشد به طوري که براي هر $d \in \mathbb{R}^n$ ، داشته باشیم:

$$\lim_{\tau \rightarrow 0} \frac{f(x + \tau d) - f(x)}{\tau} = g^T(x) d.$$

براي بردار يکه $d = e_j$ ، اگر حد

$$\lim_{\tau \rightarrow 0} \frac{f(x + \tau e_j) - f(x)}{\tau}$$

وجود داشته باشد، آن‌گاه آن را مشتق پاره‌ای f نسبت به x_j در x گویند و با نماد $\frac{\partial f(x)}{\partial x_j}$ نشان می‌دهند.

اگر مشتق پاره‌ای f در x به ازاي هر j موجود باشد، آن‌گاه گرادیان تابع f موجود است و به صورت زیر نمایش داده می‌شود:

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_j}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T.$$

تابع f را در يك ناحیه مانند D مشتق‌پذیر گویند هرگاه براي هر $x \in D$ ، $\nabla f(x)$ موجود باشد و آن را مشتق‌پذیر پیوسته یا به‌طور پیوسته مشتق‌پذیر گویند، اگر $\nabla f(x)$ تابعی پیوسته از x باشد.

تعریف ۳.۱.۱. ماتریس هسی تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ، در صورت وجود، به صورت زیر تعریف می‌شود:

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{pmatrix}.$$

تابع f را دو بار مشتق‌پذیر روی D گویند، اگر $\nabla^2 f(x)$ براي همه $x \in D$ موجود باشد و آن را دو بار مشتق‌پذیر پیوسته گویند، هرگاه $\nabla^2 f(x)$ تابعی پیوسته روی D باشد.

تعریف ۴.۱.۱. زیر مجموعه S از فضای $\mathbb{R}^n$ را محدب گویند، اگر

$$\forall x, y \in S, 0 \leq \theta \leq 1 \Rightarrow \theta x + (1 - \theta)y \in S.$$

^۱Euclidian norm

تعریف ۵.۱.۱. فرض کنید $f : \mathbb{R}^n \rightarrow \mathbb{R}$ و $S \subseteq \mathbb{R}^n$ يك مجموعه محدب باشد. تابع f را روی S محدب گویند، اگر داشته باشیم:

$$\forall x, y \in S, 0 \leq \theta \leq 1 \Rightarrow f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y).$$

ملاحظه ۱.۱.۱. در تعریف ۵.۱.۱ اگر نامساوی به ازای $\theta \in (0, 1)$ و $x \neq y$ ، اکید باشد، آن گاه f را اکیداً محدب گویند.

گزاره ۱.۱.۱. تابع به طور پیوسته مشتق پذیر f روی مجموعه S محدب است اگر و تنها اگر به ازای هر $x, y \in S$ داشته باشیم:

$$f(y) \geq f(x) + \nabla^T f(x)(y - x),$$

یا به طور معادل:

$$(\nabla f(y) - \nabla f(x))^T (y - x) \geq 0.$$

تعریف ۶.۱.۱. ماتریس $n \times n$ و متقارن H را نیمه معین مثبت گویند، اگر برای هر $x \in \mathbb{R}^n$ داشته باشیم:

$$x^T H x \geq 0.$$

تعریف ۷.۱.۱. ماتریس $n \times n$ و متقارن H را معین مثبت گویند، اگر برای هر $x \in \mathbb{R}^n$ داشته باشیم:

$$x^T H x > 0.$$

تعریف ۸.۱.۱. x^* را يك مینیمم سراسری $f : \mathbb{R}^n \rightarrow \mathbb{R}$ گویند، اگر برای هر $x \in D_f$

$$f(x) \geq f(x^*).$$

تعریف ۹.۱.۱. x^* را يك مینیمم موضعی $f : \mathbb{R}^n \rightarrow \mathbb{R}$ گویند، اگر يك همسایگی باز از x^* مانند $\mathcal{N}$ وجود داشته باشد به طوری که

$$\forall x \in \mathcal{N}; f(x) \geq f(x^*).$$

تعریف ۱۰.۱.۱. x^* را يك مینیمم موضعی اکید $f : \mathbb{R}^n \rightarrow \mathbb{R}$ گویند، اگر يك همسایگی باز از x^* مانند $\mathcal{N}$ وجود داشته باشد به طوری که

$$\forall x \in \mathcal{N}, x \neq x^*; f(x) > f(x^*).$$

گزاره ۲.۱.۱ (شرط لازم مرتبه اول). اگر x^* يك مینیمم موضعی و f در يك همسایگی باز x^* مشتق پذیر پیوسته باشد، آن گاه $\nabla f(x^*) = 0$

گزاره ۳.۱.۱ (شرط لازم مرتبه دوم). اگر x^* يك مینیمم موضعی و $\nabla^2 f$ موجود و در يك همسایگی باز x^* پیوسته باشد، آن گاه $\nabla f(x^*) = 0$ و $\nabla^2 f(x^*)$ نیمه معین مثبت است.

گزاره ۴.۱.۱ (شرط کافی مرتبه دوم). اگر $\nabla^2 f$ در يك همسایگی باز x^* پیوسته، $\nabla f(x^*) = 0$ و $\nabla^2 f(x^*)$ معین مثبت باشد، آن گاه x^* يك مینیمم موضعی اکید تابع f است.

گزاره ۵.۱.۱. اگر f تابعی محدب باشد، آن گاه هر مینیمم موضعی f ، يك مینیمم سراسری f است.

۲.۱ ساختار روش‌های عددی برای حل مسایل بهینه‌سازی

امروزه، با در نظر گرفتن مهم‌ترین قابلیت کامپیوترهای عددی که انجام عملیات تکراری به صورت کارا است، الگوریتم‌هایی تکراری برای حل مسایل بهینه‌سازی طراحی شده‌اند. در روند جستجو برای یافتن بردار $x^* \in \mathbb{R}^n$ ، به عنوان جواب بهینه برای تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ، ابتدا يك حدس اولیه x_0 در نظر گرفته می‌شود و الگوریتم موردنظر در روندی تکراری، به دنبال بهبود جواب در نقطه بعد، یعنی x_1 می‌باشد. با ادامه این روند، دنباله‌ای متوالی از نقاط اصلاح شده برای جواب حاصل می‌شود. برای مسایل بهینه‌سازی غیرخطی این دنباله در حالت کلی به جواب بهینه نمی‌رسد، اما به آن همگرا می‌شود. در حالت نامتناهی بودن دنباله تولید شده توسط الگوریتم‌های بهینه‌سازی، تخمینی از جواب بهینه (در يك همسایگی قابل قبول از آن) به عنوان جواب نهایی در نظر گرفته شده و فرآیند متوقف می‌شود. لازم به ذکر است که يك الگوریتم مناسب، نباید بیش از اندازه به زمان محاسبه یا نیازهای ذخیره‌سازی وابسته باشد و باید تا حد امکان بدون تأثیرپذیری بیش از حد از خطای داده‌ها یا خطای گرد کردن، توانایی یافتن جواب را داشته باشد. ممکن است این اهداف ناسازگار باشند؛ یعنی يك الگوریتم ممکن است برای کاهش خطا به مقدار قابل قبول، به زمان محاسبه بسیار زیادی نیاز داشته باشد. موضوعاتی از قبیل مرتبه همگرایی، نیازهای ذخیره‌سازی، سرعت و ... ابزارهای اساسی در بهینه‌سازی هستند که ارزیابی کمی و مقایسه کارکرد الگوریتم‌های مختلف را امکان‌پذیر می‌کنند. در ادامه، به تعریف مرتبه همگرایی دنباله‌ها می‌پردازیم تا بتوانیم سرعت همگرایی الگوریتم‌های بهینه‌سازی را تعریف و مقایسه کنیم.

تعریف ۱.۲.۱. فرض کنید دنباله $\{x_k\}_{k \geq 0}$ در $\mathbb{R}^n$ به x^* همگرا باشد، یعنی،

$$\lim_{k \rightarrow +\infty} \|x_k - x^*\| = 0$$

اگر عدد حقیقی $\alpha \geq 1$ و ثابت نامنفی β وجود داشته باشند به طوری که

$$\lim_{k \rightarrow +\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^\alpha} = \beta$$

آن گاه گوییم دنباله $\{x_k\}_{k \geq 0}$ دارای مرتبه همگرایی α است. به ویژه،

- هرگاه $\alpha = 1$ و $\beta \in (0, 1)$ ، دنباله به طور خطی همگرا نامیده می‌شود.
 - هرگاه $\alpha = 1$ و $\beta = 0$ ، یا $1 < \alpha < 2$ و $\beta > 0$ ، دنباله همگرای زیرخطی نامیده می‌شود.
 - هرگاه $\alpha = 2$ ، دنباله به طور مربعی همگرا نامیده می‌شود.
- توجه می‌کنیم که برای حرکت از یک نقطه به نقطه بعدی در یک روش تکراری به یک استراتژی نیاز داریم که یکی از متداول‌ترین این استراتژی‌ها، روش جستجوی خطی است که در بخش بعدی به توضیح این روش می‌پردازیم.

۱.۰.۲.۱ ساختار روش‌های مبتنی بر جستجوی خطی

در روش جستجوی خطی، الگوریتم یک جهت p_k انتخاب می‌کند و در راستای این جهت از نقطه اخیر x_k جستجو می‌کند تا به نقطه جدید

$$x_{k+1} = x_k + \alpha_k p_k$$

با مقدار تابع کمتری دست یابد. میزان حرکت در راستای این جهت، یعنی، α_k ، طول گام در تکرار k ام نام دارد. در واقع در جستجوی خطی، هدف یافتن طول گام α_k است به طوری که شرط $f(x_k + \alpha_k p_k) < f(x_k)$ برقرار باشد. محاسبه α_k به دو روش جستجوی خطی دقیق و جستجوی خطی نادقیق انجام می‌شود.

جستجوی خطی دقیق در این روش، α_k طوری انتخاب می‌شود که تابع f در امتداد p_k بیشترین کاهش را داشته باشد. به عبارت دیگر، α_k تابع ϕ متغیره $\phi(\cdot)$ با تعریف

$$\phi(\alpha) = f(x_k + \alpha p_k), \quad \alpha > 0 \quad (1.1)$$

را مینیمم می‌کند، یعنی،

$$\alpha_k = \arg \min_{\alpha > 0} f(x_k + \alpha p_k). \quad (2.1)$$

در هر تکرار یک جهت و یک طول گام جدید محاسبه می‌شود تا به نقطه جدید دست یابیم و این فرآیند به صورت تکراری ادامه می‌یابد. حل دقیق مساله (۲.۱) هزینه‌بر است. بنابراین اغلب به حل تقریبی آن بسنده می‌کنیم و از جستجوی خطی نادقیق استفاده می‌کنیم که تقریبی برای جواب مساله (۲.۱) به دست می‌دهد.

جستجوی خطی نادقیق فرض کنید که $\phi(\cdot)$ همانند (۱.۱) تعریف شود و $\bar{\alpha}_k$ کوچکترین α ی مثبتی باشد که $\phi(\alpha) = f(x_k)$. اگر $\alpha_k \approx \bar{\alpha}_k$ یا $\alpha_k \approx 0$ ، آن‌گاه تابع $\phi(\alpha)$ به ازای α_k کاهش ناچیزی خواهد داشت. بنابراین، برای آن که $\phi(\alpha)$ کاهش قابل ملاحظه‌ای داشته باشد، α_k باید از دو انتهای بازه $[0, \bar{\alpha}_k]$ فاصله داشته باشد. انجام این ویژگی در فرایندی به نام جستجوی خطی صورت می‌گیرد. در ادامه، چند روش جستجوی خطی نادقیق تشریح می‌شوند.

جستجوی خطی ولف یکی از رایج‌ترین و کاراترین روش‌های جستجوی خطی نادقیق، جستجوی خطی ولف است که در آن، α_k باید در شرایط زیر صدق کند:

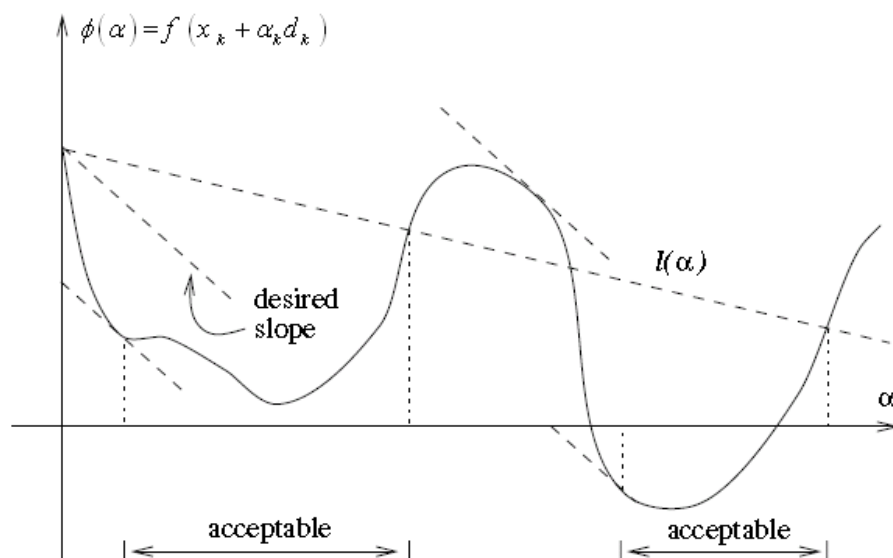
$$f(x_k + \alpha d_k) \leq f(x_k) + \rho \alpha \nabla f_k^T d_k, \quad (3.1)$$

$$\nabla f(x_k + \alpha d_k)^T d_k \geq \sigma \nabla f_k^T d_k, \quad (4.1)$$

که در آن، $0 < \rho < \sigma < 1$.

نابرابری (۳.۱) را اغلب شرط آرمیژو و نابرابری (۴.۱) را شرط انحنا می‌نامند. شرط آرمیژو، باعث دور شدن α_k از $\bar{\alpha}_k$ و شرط انحنا، باعث دور شدن α_k از ۰ می‌شود. در شکل ۱.۱ طول گام‌هایی که در شرایط ولف صدق می‌کنند، نشان داده شده‌اند که در آن،

$$l(\alpha) = f(x_k) + \rho \alpha \nabla f_k^T d_k$$



شکل ۱.۱: طول گام‌های صادق در شرایط ولف [۱۰۵]

همان‌طور که در شکل ۱.۱ مشاهده می‌شود، یک طول گام، ممکن است بدون آنکه به مینیمم کننده ϕ نزدیک باشد، در شرایط ولف صدق کند. برای رفع این مشکل، شرایط زیر برای α_k که به شرایط ولف قوی موسوم است، به کار گرفته می‌شوند:

$$f(x_k + \alpha d_k) \leq f(x_k) + \rho \alpha \nabla f_k^T d_k, \quad (5.1)$$

$$|\nabla f(x_k + \alpha d_k)^T d_k| \leq \sigma |\nabla f_k^T d_k|, \quad (6.1)$$

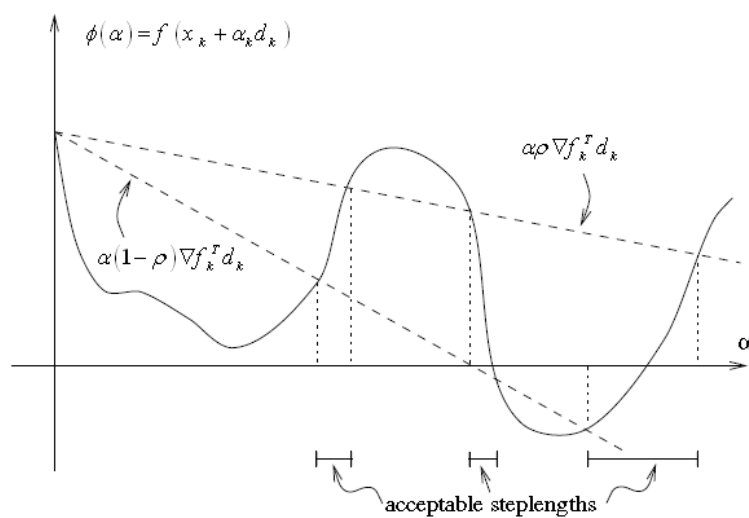
که در آن، $0 < \rho < \sigma < 1$.

جستجوی خطی گلدشتاین^۲ در این جستجوی خطی، برای دور شدن همزمان از $\bar{\alpha}_k$ و 0 ، در شرایط زیر صدق می‌کند:

$$f(x_k) + (1 - \rho) \alpha \nabla f_k^T d_k \leq f(x_k + \alpha d_k), \quad (7.1)$$

$$f(x_k + \alpha d_k) \leq f(x_k) + \rho \alpha \nabla f_k^T d_k, \quad (8.1)$$

که در آن، $0 < \rho < \frac{1}{2}$. در شکل ۲.۱ طول گام‌هایی که در شرایط گلدشتاین صدق می‌کنند، نشان داده شده‌اند. همان طور که در شکل مشاهده می‌شود، یک مشکل این روش، آنست که نابرابری (۷.۱) ممکن است همه کمینه‌کننده‌های ϕ را از بازه قابل قبول برای α_k خارج کند.



شکل ۲.۱: طول گام‌های صادق در شرایط گلدشتاین [۱۰۵]

تعریف ۲.۲.۱. بردار $p_k \in \mathbb{R}^n$ را یک جهت کاهشی برای تابع f گویند هرگاه

$$p_k^T \nabla f(x_k) < 0. \quad (9.1)$$

هم‌چنین، اگر ثابت $c > 0$ وجود داشته باشد به طوری که

$$p_k^T \nabla f(x_k) \leq -c \|\nabla f(x_k)\|^2, \quad \forall k \geq 0, \quad (10.1)$$

آن‌گاه گوییم شرط کافی کاهشی برقرار است.

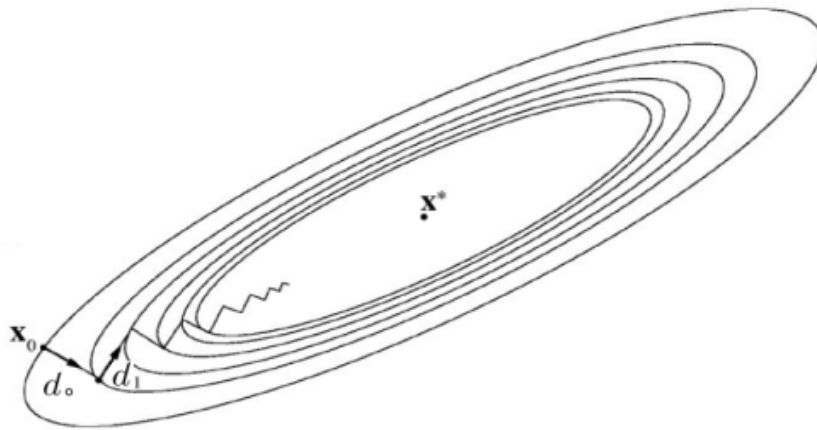
۲.۰.۲.۱ تعیین جهت جستجوی خطی

همان‌طور که اشاره کردیم، انتخاب جهت جستجو یک مساله مهم در روش جستجوی خطی است که اکنون به شرح این مساله می‌پردازیم. در این بخش به توضیح چند روش برای تعیین جهت جستجوی کاهشی برای جستجوی خطی می‌پردازیم.

^۲Goldstein

۳.۰.۲.۱ روش تندترین کاهش

یکی از جهت‌های جستجوی کاهش برای روش جستجوی خطی، جهت تندترین کاهش $-\nabla f_k$ است که از میان تمامی جهت‌ها کاهش سریع‌تری در f ایجاد می‌کند. در واقع، به راحتی می‌توان ثابت کرد که جهت $-\nabla f_k / \|\nabla f_k\|$ جهتی است که تندترین کاهش را در تابع ایجاد می‌کند. این روش از دیدگاه نظری، بسیار حائز اهمیت است؛ زیرا از ساده‌ترین روش‌هایی است که برای آن تحلیل قانع‌کننده‌ای وجود دارد. الگوریتم‌های پیشرفته‌تر، اغلب از تلاش برای اصلاح روش تندترین کاهش، به نحوی که دارای خواص همگرایی بهتری باشند، به وجود آمده‌اند. بنابراین، روش تندترین کاهش نه تنها اولین روشی است که برای حل یک مساله جدید مورد بررسی قرار می‌گیرد، بلکه به عنوان مرجعی برای سنجش عملکرد سایر روش‌ها نیز به کار می‌رود. اگرچه جهت جستجو در این روش به سادگی قابل محاسبه است اما سرعت همگرایی آن خطی است که علت این امر وجود پدیده زیگزاگ است که در شکل ۳.۱ مشخص است. علاوه بر این، سرعت همگرایی روش به شدت تحت تأثیر بد حالت بودن مساله قرار می‌گیرد.



شکل ۳.۱: روش تندترین کاهش [۱۰۵]

۴.۰.۲.۱ روش نیوتن

جهت جستجوی مهم دیگری که بسیار متداول‌تر است و از اطلاعات مرتبه دوم تابع نیز استفاده می‌کند، جهت نیوتن^۳ نام دارد که از تقریب مرتبه دوم بسط تیلور $f(x_k + p)$ حاصل می‌شود، یعنی،

$$f(x_k + p) \approx f_k + p^T \nabla f_k + \frac{1}{2} p^T \nabla^2 f_k p := m_k(p)$$

با فرض معین مثبت بودن $\nabla^2 f_k$ ، جهت نیوتن را با یافتن بردار کمینه‌کننده $m_k(p)$ به دست می‌آوریم. با مشتق‌گیری از $m_k(p)$ و مساوی صفر قرار دادن آن، به سادگی جهت زیر حاصل

^۳Newton direction

می‌شود:

$$p_k = -(\nabla^2 f_k)^{-1} \nabla f_k \quad (11.1)$$

جهت نیوتن زمانی یک جهت کاهشی است که $\nabla^2 f_k$ ماتریسی معین مثبت باشد. در واقع، اگر $\nabla^2 f_k$ معین مثبت نباشد، حتی ممکن است جهت نیوتن تعریف شده نباشد چون ممکن است $(\nabla^2 f_k)^{-1}$ وجود نداشته باشد.

روش‌هایی که از جهت نیوتن استفاده می‌کنند دارای سرعت همگرایی موضعی سریعی هستند اما محاسبه معکوس هسی در هر تکرار هزینه‌بر است. به همین علت روش‌های شبه نیوتن^۴ معرفی شدند.

۵.۰.۲.۱ روش‌های شبه نیوتن

در این روش‌ها برای یافتن جهت جستجو نیازی به محاسبه هسی نیست و به جای هسی از یک تقریب B_k استفاده می‌شود. این روش‌ها دارای سرعت همگرایی زیرخطی هستند. تکرارها براساس این موضوع بهنگام می‌شوند که تغییرات گرادیان در راستای جهت جستجو اطلاعات خوبی در مورد مشتق مرتبه دوم یعنی هسی به دست می‌دهد:

$$\nabla^2 f_k(x_{k+1} - x_k) \approx \nabla f_{k+1} - \nabla f_k \quad (12.1)$$

با قرار دادن $s_k = x_{k+1} - x_k$ و $y_k = \nabla f_{k+1} - \nabla f_k$ در رابطه (۱۲.۱) تقریب هسی در تکرار جدید یعنی B_{k+1} از رابطه

$$B_{k+1} s_k = y_k \quad (13.1)$$

به دست می‌آید. رابطه (۱۳.۱) رابطه سکانت^۵ نامیده می‌شود. با استفاده از شروط اضافی روی تقریب B_{k+1} از جمله تقارن و این مطلب که اختلاف بین B_{k+1} و B_k یک ماتریس رتبه پایین باشد، می‌توان فرمول‌هایی برای بهنگام‌سازی B_k به دست آورد. یکی از این فرمول‌ها، فرمول BFGS نام دارد که از نام افراد ارائه دهنده آن یعنی برویدن^۶، فلچر^۷، گلدفارب^۸ و شانو^۹ برگرفته شده است. فرمول BFGS به صورت

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \quad (14.1)$$

می‌باشد. توجه می‌کنیم که در این فرمول اختلاف بین B_{k+1} و B_k یک ماتریس رتبه دو است و شرط سکانت و حفظ تقارن را ارضا می‌کند. به علاوه فرمول BFGS تقریب‌هایی معین مثبت تولید می‌کند در صورتی که تقریب اولیه B_0 معین مثبت و $s_k^T y_k > 0$ باشد.

^۴Quasi-Newton methods

^۵secant

^۶Broyden

^۷Fletcher

^۸Goldfarb

^۹Shanno

هم‌چنین می‌توان از فرمولی معادل فرمول (۱۴.۱) استفاده نمود که تقریبی برای معکوس هسی به دست می‌دهد، یعنی تعریف می‌کنیم $H_k := B_k^{-1}$ و داریم:

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \quad (۱۵.۱)$$

که $\rho_k = \frac{1}{y_k^T s_k}$. در نتیجه، جهت جستجو به صورت $p_k = -H_k \nabla f_k$ حاصل می‌شود. روش BFGS را می‌توان به شکل الگوریتم ۱ بیان نمود.

الگوریتم ۱ BFGS [۱۰۵]

ورودی: نقطه اولیه x_0 ، آستانه همگرایی $\epsilon > 0$ ، تقریبی از معکوس هسی H_0 .
قرار ده $k = 0$

تازمانی که شرط گرادیان $\|\nabla f_k\| > \epsilon$ برقرار است، گام‌های زیر را تکرار کن:

- ۱: جهت جستجوی $d_k = -H_k \nabla f_k$ را محاسبه کن
- ۲: قرار ده $x_{k+1} = x_k + \alpha_k d_k$ ، که α_k با استفاده از جستجوی خطی طوری حاصل می‌شود که در شروط (۳.۱) و (۴.۱) صدق کند.
- ۳: تعریف کن $s_k = x_{k+1} - x_k$ و $y_k = \nabla f_{k+1} - \nabla f_k$
- ۴: H_{k+1} را طبق (۱۵.۱) محاسبه کن.
- ۵: قرار ده $k = k + 1$

فرضیات ۱.۲.۱. فرضیات زیر را در نظر بگیرید:

- تابع هدف f دو بار مشتق‌پذیر پیوسته باشد.
- مجموعه سطح $L = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ محدب باشد و اعداد ثابت مثبت m و M وجود داشته باشند به طوری که

$$m\|z\|^2 \leq z^T G(x) z \leq M\|z\|^2 \quad (۱۶.۱)$$

برای هر $x \in L, z \in \mathbb{R}^n$ که در آن G میانگین هسی است، یعنی،

$$G(x) = \int_0^1 \nabla^2 f(x + \tau \alpha p) d\tau.$$

بخش دوم این فرضیات نتیجه می‌دهد که $G(x)$ روی L معین مثبت است و f یک کمینه‌کننده یکتا x^* در L دارد.

قضیه ۱.۲.۱. [۱۰۵] فرض کنید H_0 ماتریس متقارن معین مثبت اولیه باشد و x_0 نقطه اولیه‌ای باشد که در فرضیات ۱.۲.۱ صدق می‌کند. در این صورت دنباله $\{x_k\}$ تولید شده توسط الگوریتم ۱ به کمینه‌کننده x^* برای تابع f همگرا می‌شود.

۶.۰.۲.۱ روش BFGS حافظه محدود

در این بخش، به شرح مختصری از روش BFGS حافظه محدود (LBFGS)^{۱۰} می‌پردازیم. هر مرحله از روش BFGS یک دنباله جواب به صورت $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k H_k \nabla f_k$ تولید می‌کند. α_k طول گام و H_k تقریب معکوس هسی است که در هر تکرار به وسیله فرمول

$$H_{k+1} = V_k^T H_k V_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T, \quad (۱۷.۱)$$

به‌نگام می‌شود، به‌طوری که در آن،

$$\rho_k = \frac{1}{\mathbf{y}_k^T \mathbf{s}_k}, \quad V_k = I - \rho_k \mathbf{y}_k \mathbf{s}_k^T,$$

و

$$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \quad \mathbf{y}_k = \nabla f_{k+1} - \nabla f_k. \quad (۱۸.۱)$$

از آنجا که تقریب معکوس هسی H_k به‌طور کلی متراکم است، هزینه‌ی ذخیره و کار با آن زمانی که تعداد متغیرها زیاد است به شدت افزایش می‌یابد. برای جلوگیری از این مشکل یک نسخه اصلاح شده از H_k را با ذخیره کردن یک تعداد معینی (M) از جفت بردارهای $\{\mathbf{s}_i, \mathbf{y}_i\}$ بدست می‌آوریم. یعنی زمانی که $k > M$ قدیمی‌ترین جفت بردار $\{\mathbf{s}_i, \mathbf{y}_i\}$ با جفت بردار جدید $\{\mathbf{s}_k, \mathbf{y}_k\}$ از گام جاری جایگزین می‌شود. به این ترتیب، مجموعه جفت بردارها شامل اطلاعات انحنای M تکرار اخیر است. سپس می‌توانیم در تکرارهای بعدی حاصل ضرب $H_k \nabla f_k$ را با استفاده از جفت بردارهای $\{\mathbf{s}_i, \mathbf{y}_i\}$ و ∇f_k بدست آوریم. تجربه‌های عملی نشان داده است که مقادیر M بین ۳ تا ۲۰ اغلب مناسب هستند و نتایج قابل قبولی ارائه می‌کنند.

حال فرآیند به‌به‌نگام‌سازی را شرح می‌دهیم. در تکرار k ام، نقطه جاری $\mathbf{x}_k$ است و جفت بردارهای $\{\mathbf{s}_i, \mathbf{y}_i\}$ برای هر $i = k - M, \dots, k - 1$ نشان داده می‌شود. ابتدا تقریب هسی اولیه H_k^0 را محاسبه می‌کنیم (در مقایسه با الگوریتم BFGS استاندارد این تقریب اولیه در هر تکرار می‌تواند تغییر کند). سپس با استفاده مکرر از فرمول (۱۷.۱) تقریب H_k به صورت زیر مشخص می‌شود.

$$\begin{aligned} H_k &= (V_{k-1}^T \dots V_{k-M}^T) H_k^0 (V_{k-M} \dots V_{k-1}) \\ &+ \rho_{k-M} (V_{k-1}^T \dots V_{k-M+1}^T) \mathbf{s}_{k-M} \mathbf{s}_{k-M}^T (V_{k-M+1} \dots V_{k-1}) \\ &+ \rho_{k-M+1} (V_{k-1}^T \dots V_{k-M+2}^T) \mathbf{s}_{k-M+1} \mathbf{s}_{k-M+1}^T (V_{k-M+2} \dots V_{k-1}) \\ &+ \dots \\ &+ \rho_{k-1} \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T. \end{aligned}$$

با توجه به این عبارت می‌توانیم با یک روش بازگشتی حاصل ضرب $H_k \nabla f_k$ را بدست آوریم.

^{۱۰}Limited memory BFGS

الگوریتم ۲ الگوریتم دو حلقه بازگشتی LBFGS [۱۰۵]

۱: قرار دهید $\mathbf{q} \leftarrow \nabla f_k$ ۲: برای $i = k - 1, k - 2, \dots, k - M$ قرار دهید:

$$\alpha_i \leftarrow \rho_i \mathbf{s}_i^T \mathbf{q},$$

$$\mathbf{q} \leftarrow \mathbf{q} - \alpha_i \mathbf{y}_i.$$

پایان حلقه

۳: قرار دهید: $\mathbf{r} \leftarrow H_k^\circ \mathbf{q}$ ۴: برای $i = k - M, k - M + 1, \dots, k - 1$ قرار دهید:

$$\beta \leftarrow \rho_i \mathbf{y}_i^T \mathbf{r},$$

$$\mathbf{r} \leftarrow \mathbf{r} + \mathbf{s}_i (\alpha_i - \beta).$$

پایان حلقه

۵: نتیجه $\mathbf{r} = H_k \nabla f_k$

یک روش برای انتخاب H_k° که کارایی آن در عمل اثبات شده است انتخاب $H_k^\circ = \gamma_k I$ است به‌طوری که:

$$\gamma_k = \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T \mathbf{y}_{k-1}}. \quad (19.1)$$

این استراتژی در عمل به خوبی کار می‌کند. در واقع هنوز ثابت نشده است که استراتژی دیگری همواره بهتر از این استراتژی عمل می‌کند. در $M - 1$ تکرار اول اگر ماتریس اولیه $H_\circ$ در هر دو روش یکسان باشد و الگوریتم BFGS حافظه محدود در هر تکرار برابر $H_\circ^k = H_\circ$ باشد، الگوریتم ۳ با الگوریتم BFGS استاندارد معادل است. ضعف اصلی روش LBFGS این است که روی مسایل بدحالت به کندی همگرا است.

۷.۰.۲.۱ روش بی‌حافظه BFGS

خطر مهمی که می‌تواند کارایی عددی روش‌های شبه نیوتن را تهدید کند، به بد حالت بودن تقریب‌های ماتریسی تولید شده در این روش‌ها برمی‌گردد که منجر به ناپایداری عددی روش‌های شبه نیوتن خواهد شد. لذا، برای بدست آوردن توزیع مناسب مقادیر ویژه و نیز بهبود عدد حالت تقریب‌های بدست آمده برای معکوس ماتریس هسی و در نتیجه افزایش پایداری عددی در روش‌های تکراری مورد نظر، روش‌های شبه نیوتن مقیاس‌بندی شده پیشنهاد شده‌اند. در فرمول بهنگام‌سازی روش‌های شبه نیوتن مقیاس‌بندی شده به جای H_k از $\theta_k H_k$ استفاده می‌شود که در آن $\theta_k \in \mathbb{R}$ پارامتر مقیاس نامیده می‌شود. به عنوان نمونه، فرمول تکراری

الگوریتم ۳ الگوریتم BFGS حافظه محدود [۱۰۵]

پارامترهای ورودی: نقطه اولیه $\mathbf{x}_0$ و عدد صحیح M انتخاب کرده و قرار دهید $k = 0$.
تازمانی که شرط توقف برقرار نیست گام‌های زیر را تکرار کنید:

- ۱: H_k° را با استفاده از فرمول $H_k^\circ = \gamma_k I$ محاسبه کنید، به طوری که $\gamma_k = \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T \mathbf{y}_{k-1}}$.
- ۲: با استفاده از الگوریتم ۲ حاصل ضرب هسی در بردار $H_k \nabla f_k$ را محاسبه کنید و قرار دهید $\mathbf{p}_k = -H_k \nabla f_k$.
- ۳: نقطه جدید $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$ را محاسبه کنید به طوری که α_k در شرایط ولف صدق کند.
- ۴: اگر $k > M$ ، جفت بردار $\{\mathbf{s}_{k-M}, \mathbf{y}_{k-M}\}$ را از حافظه حذف کنید و $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ و $\mathbf{y}_k = \nabla f_{k+1} - \nabla f_k$ را محاسبه و ذخیره کنید. سپس قرار دهید $k = k + 1$.

بهنگام‌سازی مقیاس‌بندی شده BFGS به صورت زیر است:

$$\hat{H}_{k+1}^{BFGS} = \theta_k H_k - \theta_k \frac{\mathbf{s}_k \mathbf{y}_k^T H_k + H_k \mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} + \left(1 + \theta_k \frac{\mathbf{y}_k^T H_k \mathbf{y}_k}{\mathbf{s}_k^T \mathbf{y}_k}\right) \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}. \quad (20.1)$$

کارترین انتخاب‌ها برای پارامتر مقیاس θ_k در بهنگام‌سازی فوق که روش‌های شبه نیوتن خودمقیاس را به دست می‌دهند، توسط اورن^{۱۱} و اسپدیکاتو^{۱۲} و اورن و لئونبرگر^{۱۳} معرفی شده‌اند [۱۰۷، ۱۰۸]:

$$\theta_k^{OL} = \frac{\mathbf{s}_k^T H_k^{-1} \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{y}_k} \quad (21.1)$$

و

$$\theta_k^{OS} = \frac{\mathbf{s}_k^T \mathbf{y}_k}{\mathbf{y}_k^T H_k \mathbf{y}_k}. \quad (22.1)$$

با وجود کارایی عددی مطلوب روش‌های شبه نیوتن مقیاس‌بندی شده، این روش‌ها برای حل مسایل بزرگ مقیاس مناسب نیستند، زیرا به ذخیره‌سازی و انجام اعمال حسابی با ماتریس H_k نیاز دارند که این امر در ابعاد بالا بسیار پرهزینه است. در جهت رفع این مشکل، روش‌های شبه نیوتن مقیاس‌بندی شده بی حافظه BFGS^{۱۴} مطرح شدند که در آن‌ها به جای H_k از ماتریس همانی در فرمول‌های بهنگام‌سازی استفاده می‌شود. به عنوان نمونه، فرمول بهنگام‌سازی

^{۱۱}Oren

^{۱۲}Spedicato

^{۱۳}Luenberge

^{۱۴}memoryless BFGS

مقیاس بندی شده بی‌حافظه BFGS را به صورت زیر داریم:

$$\hat{H}_{k+1}^{BFGS} = \theta_k I - \theta_k \frac{s_k y_k^T + y_k s_k^T}{s_k^T y_k} + \left(1 + \theta_k \frac{y_k^T y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k}. \quad (23.1)$$

هم‌چنین، نسخه‌های بی‌حافظه پارامترهای مقیاس (۲۱.۱) و (۲۲.۱) به ترتیب به صورت زیر هستند [۱۴]:

$$\theta_k^{OL} = \frac{\|s_k\|^2}{s_k^T y_k} \quad (24.1)$$

یا

$$\theta_k^{OS} = \frac{s_k^T y_k}{\|y_k\|^2}. \quad (25.1)$$

لازم به ذکر است که جهت‌های جستجو در روش‌های مقیاس بندی شده بی‌حافظه BFGS با تعداد محدودی ذخیره‌سازی برداری و انجام تعداد اندکی ضرب داخلی از فرمول زیر قابل محاسبه است:

$$d_{k+1}^{BFGS} = -H_{k+1}^{BFGS} g_{k+1} = -\theta_k g_{k+1} + \theta_k \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k + \theta_k \frac{s_k^T g_{k+1}}{s_k^T y_k} s_k - \left(1 + \theta_k \frac{y_k^T y_k}{s_k^T y_k}\right) \frac{s_k^T g_{k+1}}{s_k^T y_k} s_k$$

ضمناً، اگر تابع هدف به طور یکنواخت محدب باشد، جهت‌های تولید شده توسط این روش در شرط کافی کاهشی نیز صدق می‌کند [۱۳، ۱۵].

۸.۰.۲.۱ روش گاوس- نیوتن

روش گاوس- نیوتن از روش‌های اساسی برای حل مسایل کمترین مربعات غیر خطی است. در این مسایل تابع هدف به صورت زیر است:

$$f(x) = \frac{1}{2} \sum_{j=1}^m \varphi_j^2(x) \quad (26.1)$$

که در آن هر φ_j یک تابع هموار از $\mathbb{R}^n$ به $\mathbb{R}$ است. هر φ_j را یک باقیمانده می‌گویند و فرض می‌شود $m \geq n$.

با در نظر گرفتن $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ که $\varphi(x) = (\varphi_1(x), \varphi_2(x), \dots, \varphi_m(x))^T$ تابع f را می‌توان به صورت $f(x) = \frac{1}{2} \|\varphi(x)\|^2$ نوشت.

گرادیان و هسی تابع f به صورت زیر بدست می‌آید:

$$\nabla f(x) = \sum_{j=1}^m \varphi_j(x) \nabla \varphi_j(x) = J(x)^T \varphi(x)$$

$$\begin{aligned} \nabla^2 f(x) &= \sum_{j=1}^m \nabla \varphi_j(x) \nabla \varphi_j(x)^T + \sum_{j=1}^m \varphi_j(x) \nabla^2 \varphi_j(x) \\ &= J(x)^T J(x) + \sum_{j=1}^m \varphi_j(x) \nabla^2 \varphi_j(x) \end{aligned} \quad (27.1)$$

که در آن $J(x)$ ماتریس ژاکوبی و به صورت زیر است:

$$J(x) = \left[\frac{\partial \varphi_j}{\partial x_i} \right]_{\substack{j=1, \dots, m \\ i=1, \dots, n}} = \begin{bmatrix} \nabla \varphi_1(x)^T \\ \nabla \varphi_2(x)^T \\ \vdots \\ \nabla \varphi_m(x)^T \end{bmatrix}$$

در رابطه (۲۷.۱)، $J(x)^T J(x)$ جمله اول هسی و $\sum_{j=1}^m \varphi_j(x) \nabla^2 \varphi_j(x)$ جمله دوم هسی نامیده می‌شود و از آن جاکه φ_j ها عموماً معرف خطا هستند جمله اول هسی مهم‌تر از جمله دوم هسی می‌باشد.

روش گاوس – نیوتن یک روش نیوتن اصلاح شده است که در آن به جای حل معادلات نیوتن استاندارد، $\nabla^2 f(x_k) d = -\nabla f(x_k)$ ، دستگاه زیر برای بدست آوردن جهت جستجو d_k^{GN} حل می‌شود:

$$J_k^T J_k d = -J_k^T \varphi_k$$

که در آن $J_k = J(x_k)$ و $\varphi_k = \varphi(x_k)$ است.

روش گاوس – نیوتن دارای سه خاصیت زیر است:

- (۱) فقط به مشتقات مرتبه اول نیاز دارد و نیازی به محاسبه هسی باقی‌مانده‌ها نیست.
 - (۲) به طور ذاتی ماتریس $J^T J$ نیمه معین مثبت است و لذا جهت‌های بدست آمده از این روش صعودی نمی‌باشند.
 - (۳) تقریب هسی برای φ_j های خطی یا باقی‌مانده صفر دقیق است.
- طبق نتایج تئوری از میر^{۱۵} [۹۱]، دنباله $\{x_k\}$ تولید شده به روش گاوس – نیوتن در نامساوی زیر صدق می‌کند:

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq \frac{\tau}{\mu} = \rho$$

که τ بزرگترین مقدار ویژه جمله دوم هسی و μ کمترین مقدار ویژه جمله اول هسی است و این دنباله برای $\rho < 1$ به همگراست (این در صورتی است که یا باقی‌مانده در جواب کوچک باشد یا تابع باقی‌مانده خطی باشد).

لذا وقتی که $J_k^T J_k$ یک تقریب خوب از $\nabla^2 f$ باشد نرخ همگرایی روش گاوس – نیوتن مشابه روش نیوتن است، اما وقتی J_k رتبه ناقص یا نزدیک رتبه ناقص باشد روش گاوس – نیوتن با مشکل مواجه می‌شود که با بکار بردن الگوریتم **لونیبرگ – مارکوارت**^{۱۶} [۱۰۵] از این مشکل دوری می‌شود. همچنین این روش برای مسایل غیرخطی یا باقی‌مانده بزرگ دارای کارایی پایین‌تری نسبت به روش‌های نیوتن و یا شبه نیوتن است. با این وجود کارایی هر دو

^{۱۵} Meyer

^{۱۶} Levenberg–Marquardt

روش نیوتن و شبه نیوتن قبل از دستیابی به همسایگی جواب در تکرارهای آغازین پایین‌تر از روش‌های گاوس – نیوتن و لونبرگ – مارکوارت می‌باشند. از طرف دیگر اغلب نمی‌دانیم که آیا مساله، باقی‌مانده بزرگ در جواب تولید می‌کند یا کوچک. به همین خاطر الگوریتم‌های ترکیبی در نظر گرفته می‌شوند که اگر مساله باقی‌مانده کوچک تولید کند رفتار مشابه گاوس – نیوتن یا لونبرگ – مارکوارت داشته باشند اما اگر باقی‌مانده در جواب بزرگ باشد مشابه گام‌های روش نیوتن یا شبه نیوتن رفتار کنند [۱۰۵].

فصل ۲

مقدمه‌ای بر مفاهیم مکان‌یابی

مسایل مکان‌یابی مسایلی هستند که وقتی به دنبال یافتن محل و نحوه استقرار بهینه یک یا چند سرویس‌دهنده بر اساس عوامل و متغیرهای مؤثر بر مکان‌یابی هستیم، با آن‌ها مواجه می‌شویم. معمولاً در این مسایل فرض بر این است که تعدادی مشتری (سرویس‌گیرنده یا نقطه تقاضا) موجودند و می‌خواهیم بهترین مکان برای استقرار سرویس‌دهندگان جدید را به گونه‌ای بیابیم که هزینه حمل و نقل، تأثیرات محیطی نامطلوب، زمان سرویس‌دهی، سود، کیفیت، نحوه توزیع سرویس و رعایت انصاف و برابری در آن، کسب بیشترین سهم در بازار و غیره، بسته به شرایط مساله بهینه شوند.

مکان‌یابی تأسیسات و تسهیلات، از کاربردی‌ترین زمینه‌های تحقیقاتی می‌باشد که در بسیاری از شاخه‌های علوم گسترده شده است. از جمله این شاخه‌ها می‌توان به تحقیق در عملیات، علوم مدیریت و تخصیص منابع، مهندسی صنایع، جغرافیا، اقتصاد، علوم کامپیوتر، ریاضیات، بازاریابی، مهندسی برق، برنامه‌ریزی شهری، علوم نظامی و دیگر شاخه‌های مرتبط اشاره کرد.

از کاربردهای معروف مسایل مکان‌یابی می‌توان مکان‌یابی انبارها، کارخانه‌ها، ماشین‌آلات کارخانه، بیمارستان‌ها، نمایندگی‌های فروش، آتش‌نشانی‌ها، پاسگاه‌های پلیس، پایانه‌های مسافربری، اجزاء الکترونیکی در مدارها، آذیر خطر، چاه‌های نفت، تأسیسات نظامی، نیروگاه‌ها، آزمایشگاه‌ها و غیره را نام برد.

۱.۲ زمینه‌های عمده تحقیق در نظریه مکان‌یابی

شاید بتوان پیدایش مساله مکان‌یابی را به زمانی نسبت داد که در قرن هفدهم میلادی فرما^۱ مساله‌ای را به این صورت مطرح نمود: ”فرض کنید سه نقطه در صفحه داده شده است، نقطه چهارم را به گونه‌ای بیابید که مجموع فاصله‌های آن تا سه نقطه داده شده کمینه گردد.“^۲ **توریچلی** در سال ۱۶۴۰ مشاهده کرد که اگر سه نقطه موجود را سه رأس یک مثلث در نظر بگیریم، محل برخورد دایره‌های محیطی مثلث‌های متساوی الاضلاعی که بر روی اضلاع مثلث اصلی و رو به بیرون آن ساخته می‌شوند، جواب مساله است. بدین دلیل، نقطه بهینه را نقطه توریچلی می‌نامند. در سال ۱۶۴۷ **کاوالیری**^۳ نشان داد که زاویه بین نقطه توریچلی و سه نقطه مورد نظر، ۱۲۰ درجه است. **سیمپسون**^۴ در سال ۱۷۵۰ ثابت کرد که اگر یک مثلث متساوی الاضلاع بر سه نقطه داده شده گذرانده شود و رئوس این مثلث به یکی از این سه نقطه که بر روی ضلع مقابل قرار گرفته وصل شود، محل تقاطع این سه خط، نقطه توریچلی را مشخص می‌کند. از این رو این سه خط را خطوط سیمپسون می‌نامند.

حالت استثنایی که تا سال ۱۸۳۴ بدون حل باقی ماند این بود که سه نقطه داده شده مثلثی را تشکیل دهند که یک رأس زاویه بیشتر یا مساوی ۱۲۰ درجه داشته باشد، که **هین**^۵ ثابت کرد همان رأسی که دارای زاویه بیشتر یا مساوی ۱۲۰ درجه است، جواب بهینه مساله می‌باشد. او همچنین ثابت کرد طول خطوط سیمپسون با هم مساوی هستند و طول آن‌ها با فاصله نقطه بهینه از سه نقطه داده شده نیز برابر است. **فسبندر**^۶ مفهوم دوگانی را برای مساله فرما مطرح کرد. بعدها در سال ۱۹۶۷ **کوهن**^۷ به‌طور مبسوطی مفهوم دوگانی را مورد بررسی قرار داد و نتایج جالبی از تحقیقات وی منتشر گردید [۸۴].

نظریه مکان‌یابی به شکلی که امروزه مورد استفاده قرار می‌گیرد، با منتشر شدن کتاب **آلفرد وبر**^۸ در سال ۱۹۰۹ تحت عنوان *Über den Standort der Industrien* متولد شد [۱۳۷، ۲]. وبر در این کتاب نتایج تحقیقات خود را در مورد صنایع کارخانه‌ای ارائه کرد. در واقع می‌توان گفت که اولین تعریف مساله مکان‌یابی به‌صورت کاربردی توسط وبر ارائه شده است [۱۳۷]. نظریه وبر دارای سه اصل اساسی است. اول اینکه محل مواد اولیه معلوم است. دوم اینکه مکان و میزان مصرف مشتریان مشخص است. تعداد این مکان‌ها متناهی است و جدا از هم هستند. هر تولید کننده بازار نامحدودی در اختیار دارد و امتیازات انحصاری از نظر انتخاب مکان وجود ندارد. سوم اینکه چندین مکان ثابت برای عرضه نیروی کار وجود دارد به‌طوری که

^۱ Fermat^۲ Torricelli^۳ Cavalieri^۴ Simpson^۵ Heinen^۶ Fesbender^۷ Kuhn^۸ Alfred Weber

این نیروی کار دارای قابلیت تحرک و جابجایی نبوده و عرضه آن نامحدود است. اما مطالعات جدی بر روی مکان‌یابی از زمانی شروع شد که در سال ۱۹۶۴ **حکیمی^۹**، تابع هدف را به دو صورت **کمترین مجموع^{۱۰}** و **مینیماکس^{۱۱}** مطرح کرد [۶۳]. او همچنین مطالعاتی بر روی مسایل مکان‌یابی روی شبکه‌ها انجام داد. در واقع حکیمی کسی بود که روابط p -میان را ابداع و مساله p -میان را معرفی نمود. از آن زمان به بعد مطالعات زیادی بر روی مسایل مکان‌یابی انجام گرفته است.

معمولاً در مسایل مکان‌یابی سرویس‌دهندگانی که می‌خواهیم مکان استقرار آن‌ها را بیابیم، مطلوب در نظر گرفته می‌شوند و تنها تأثیرات مثبت آن‌ها را به حساب می‌آوریم. بنابراین نزدیک بودن به آن‌ها نوعی جاذبه و دور بودن از آن‌ها نوعی دافعه تلقی می‌شود. در جوامع مدرن و پیشرفته امروزی با فعالیت‌ها و سرویس‌دهندگانی روبه‌رو هستیم که علاوه بر ارائه خدمات، تأثیرات منفی و مخربی نیز بر محیط پیرامون خود دارند. کارگاه‌های تولیدی، مثال آشکار این مورد هستند. کارگاه کالای مورد نیاز جامعه را تولید کرده و از طرف دیگر آلاینده محیط زیست است. مراکز ورزشی در عین حال که باید در دسترس باشند و امکانات ورزشی و تفریحی مناسب را ارائه دهند، می‌توانند تأثیرات منفی همچون سر و صدا، ازدحام جمعیت و یا مشکلاتی از این قبیل را نیز داشته باشند.

مطالعه مدل‌هایی که هر دو ویژگی را به‌طور همزمان در خود دارند، انطباق بیشتری با مسایل واقعی و کاربردی دارد و به سرعت در حال پیشرفت است. **گلدمن و دیرینگ^{۱۲}** [۵۹] و **چرچ و گارفینکل^{۱۳}** [۳۲] اولین کسانی بودند که مواردی را پیشنهاد کردند که بعدها مدل‌های مکان‌یابی مضر^{۱۴} نام گرفتند. در این دسته مدل‌ها، مشتریان فرض مطلوب بودن سرویس‌دهنده‌ها را کنار می‌گذارند و دیگر سعی در نزدیک‌تر بودن به آن‌ها را نداشته و تا حدی از آن‌ها دوری می‌کنند. به عنوان کاربردی از این مدل‌ها می‌توان به مکان‌یابی راکتورهای هسته‌ای، مراکز تصفیه آب و اماکن دفن زباله اشاره کرد. در حقیقت ویژگی‌های این نوع سرویس‌دهندگان نوعی دافعه ایجاد می‌کند و سبب می‌شود که مشتریان آن‌ها را از خود دور کنند.

می‌توان گفت اولین طبقه‌بندی مدل‌های مختلف مکان‌یابی توسط **هندلر^{۱۵}** و **میرچندانی^{۱۶}** ارائه شد [۶۵]. پس از آن طبقه‌بندی‌های دیگری از جمله توسط **هاماچر^{۱۷}** و **نیکل^{۱۸}** انجام شد [۶۴]. همچنین بررسی‌هایی از کارهای انجام شده در این زمینه ارائه شد که از آن جمله

^۹Hakimi

^{۱۰}Minisum

^{۱۱}Minimax

^{۱۲}Golman and Dearing

^{۱۳}Church and Garfinkel

^{۱۴}Obnoxious

^{۱۵}Handler

^{۱۶}Mirchandani

^{۱۷}Hamacher

^{۱۸}Nickel

تنسل^{۱۹} و همکاران [۱۲۷]، کرارپ^{۲۰} و پروزن^{۲۱} [۸۳]، هانسن^{۲۲} و همکاران [۶۸]، اون^{۲۳} و دسکین^{۲۴} [۱۰۹]، کاپانرا^{۲۵} [۳۱]، اسکاپارا^{۲۶} و اسکاتلا^{۲۷} [۱۱۸] لیستی از کارهای انجام شده در زمینه‌های مختلف مکان‌یابی را ارائه کرده‌اند. کارنت^{۲۸} و همکاران نیز کاربردهای مسایل مکان‌یابی را بیان نموده‌اند [۳۷].

به‌طور ویژه در دو دهه اخیر نیز مطالعات گوناگونی بر روی مسایل مکان‌یابی صورت گرفته است که به‌عنوان نمونه کامپیل^{۲۹} مساله مکان‌یابی گسسته هاب را در سال ۱۹۹۴ مورد مطالعه قرار داد [۳۰]. در سال ۲۰۰۱ کلامروس^{۳۰} به مساله تک وسیله‌ای با موانع چند وجهی پرداخت و یک الگوریتم قطعی و یک راه حل ابتکاری برای آن ارائه کرد [۸۰]. براتکورن^{۳۱} و همکاران در سال ۲۰۰۲ الگوریتم هوشمند سریعی برای رده وسیعی از مسایل مکان‌یابی پوشش ارائه کردند [۲۶]. همچنین یک رویکرد شاخه و قیمت در سال ۲۰۰۵ توسط سنی^{۳۲} و همکاران ارائه شد [۱۲۰]. زانگ^{۳۳} در سال ۲۰۰۶ الگوریتمی شبه حریصانه برای مساله مکان‌یابی دو سطحی پیشنهاد کرد [۱۴۵]. در سال ۲۰۰۷ ساهین^{۳۴} و سورات^{۳۵} مدل‌های گوناگون مساله مکان‌یابی سلسله‌مراتبی را مورد بررسی قرار دادند [۱۱۷]. بالسیک^{۳۶} و بیامون^{۳۷} در سال ۲۰۰۸ به مساله مکان‌یابی برای مساله کمک‌های بشر دوستانه پرداختند [۱۶]. در سال ۲۰۰۹ جمالیان و فتحعلی به مساله مکان‌یابی با کمترین مقدار خطای مطلق پرداختند [۷۵]. در سال ۲۰۱۱ کالکسیکس^{۳۸} مساله میانه چند وسیله‌ای با وزن‌های مثبت و منفی روی گراف را مورد مطالعه قرار داد [۷۷]. فرهانی و همکاران مساله مکان‌یابی پوشش را در سال ۲۰۱۲ مورد مطالعه قرار دادند. در سال ۲۰۱۳ لی^{۳۹} به مساله مکان‌یابی بدون ظرفیت پرداخت و الگوریتمی تقریبی برای حل آن ارائه نمود [۸۵]. در همین سال برسنو^{۴۰} الگوریتم شاخه و

^{۱۹}Tansel

^{۲۰}Krarp

^{۲۱}Pruzan

^{۲۲}Hansen

^{۲۳}Owen

^{۲۴}Deskin

^{۲۵}Cappanera

^{۲۶}Scaparra

^{۲۷}Scutella

^{۲۸}Current

^{۲۹}Campbell

^{۳۰}Klamroth

^{۳۱}Brotcorne

^{۳۲}Senne

^{۳۳}Zhang

^{۳۴}Sahin

^{۳۵}Sural

^{۳۶}Balcik

^{۳۷}Beamon

^{۳۸}Kalcsics

^{۳۹}Li

^{۴۰}Beresnev

کران را برای مساله مکان‌یابی رقابتی ارائه نمود [۲۰]. در سال ۲۰۱۴ رحمانی و میرحسینی الگوریتم ترکیبی ژنتیک کرم شب تاب را برای مساله مکان‌یابی با ظرفیت معرفی کردند [۱۱۳]. هو^{۴۱} در سال ۲۰۱۵ به مساله با ظرفیت تک منبع پرداخته و یک الگوریتم ابتکاری تکراری تابو را برای آن ارائه کرد [۷۰]. در سال ۲۰۱۶ مانتویا^{۴۲} و همکاران مساله مکان‌یابی با ظرفیت چند محصوله را مورد مطالعه قرار دادند [۹۶]. در حال حاضر به دلیل افزایش روز افزون کاربردهای مسایل مکان‌یابی در زندگی روزمره، بسیاری از محققین و پژوهشگران به مطالعه در این شاخه از بهینه‌سازی پرداخته‌اند [۱۲۹، ۴۲، ۱۱۲].

مسایل مکان‌یابی در توابع هدف، فواصل، تعداد و اندازه سرویس‌دهنده‌هایی که باید تاسیس شوند و در چندین عامل دیگر متفاوت هستند. از این‌رو بسته به نوع هدف و با استفاده از هر یک از این شاخص‌های مختلف به مدل‌های مختلف مساله مکان‌یابی دست خواهیم یافت.

۲.۲ مولفه‌های اساسی مسایل مکان‌یابی

در حالت کلی یک مساله مکان‌یابی دارای عناصر و مولفه‌های اساسی است که در ادامه به آن‌ها اشاره خواهد شد.

۱.۲.۲ فضای مسایل مکان‌یابی

معمولاً فضای مسایل مکان‌یابی یکی از سه دسته زیر است:

• فضای پیوسته

هرگاه فضای مکان تأسیسات و نقاط تقاضا پیوسته باشند، یعنی توسط متغیرهایی که به صورت پیوسته تغییر می‌کنند، نظیر مختصات مشخص شود، مدل را پیوسته می‌گویند. در این مدل‌ها مکان سرویس‌دهنده را باید در فضای d -بعدی $\mathbb{R}^d$ پیدا نمود. اکثر مسایل مکان‌یابی پیوسته در فضای حداقل دو بعدی تعریف می‌شوند. مسایل دو بعدی به خاطر ویژگی‌های هندسی و قابلیت ادراک بصری آن‌ها بیشتر مورد توجه قرار می‌گیرند. در مواجهه با مسایل مکان‌یابی پیوسته از ابزارهای هندسه و آنالیز استفاده می‌شود.

• فضای گسسته

مدل‌های مکان‌یابی گسسته مدل‌هایی هستند که در آن‌ها باید مکان سرویس‌دهندگان تنها روی نقاط از پیش تعیین شده پیدا شوند. در این مدل‌ها مجموعه‌ای متناهی از نقاط کاندیدا داده شده است که توسط متغیرهای گسسته نشان داده می‌شوند. در مدل‌های مکان‌یابی

^{۴۱}Ho

^{۴۲}Montoya

پیوسته بر خلاف مدل‌های گسسته نمی‌توان لیست جامع و کاملی از نقاط در دسترس را ارائه داد. از این جهت مدل‌های پیوسته را مدل‌های ایجاد مکان و مدل‌های گسسته را مدل‌های انتخاب مکان نیز خوانده‌اند.

• فضای شبکه‌ای

به‌طور کلی مسایل مکان‌یابی بر روی فضاهای شبکه‌ای رده‌ای دیگر از مسایل مکان‌یابی گسسته است که به جستجو برای یافتن بهترین مکان برای تاسیس سرویس‌دهنده‌ها روی شبکه می‌پردازد.

یعنی در این رده از مسایل، هدف پیدا کردن مجموعه‌ای از نقاط روی شبکه مفروض است به‌طوری که هزینه کل سرویس‌دهی این نقاط (سرویس‌دهنده‌ها) تا نقاط تقاضا (مشری‌ها) کمینه شود.

در حوزه مکان‌یابی با فضای شبکه‌ای، مسایلی وجود دارند که در حالت کلی مسایلی NP - سخت هستند. از جمله این مسایل می‌توان مساله مکان‌یابی p - میانه، مساله مکان‌یابی p - مرکز، مساله مکان‌یابی بدون ظرفیت و مساله مکان‌یابی پوشش را نام برد. در سال ۱۹۶۸ تیتز^{۴۳} و بارت^{۴۴} به حل مساله p - میانه روی گراف پرداختند و روشی جدید برای یافتن p - میانه راسی ارائه کردند [۱۲۸]. در سال ۱۹۹۱ کو^{۴۵} و لی^{۴۶} الگوریتمی با پیچیدگی زمانی $O(n)$ برای مساله وزن‌دار ۲ - مرکز و ۳ - مرکز با فاصله مستطیلی ارائه کردند [۸۱]. کاروین^{۴۷} و حکیمی در سال ۱۹۷۹ نشان دادند که مساله مکان‌یابی p - میانه روی شبکه مساله‌ای NP - سخت است و الگوریتمی ارائه کردند که p - میانه درخت را می‌یابد [۷۸]. همچنین آن‌ها در سال ۱۹۷۹ الگوریتمی برای یافتن p - مرکز روی شبکه ارائه نمودند [۷۹]. در سال ۱۹۹۴ کامپیل در مقاله‌ای تحت عنوان "فرمول‌بندی عدد صحیح مساله مکان‌یابی گسسته هاب"، فرمول‌بندی عدد صحیح را برای چهار نوع از مسایل هاب شامل مساله p - میانه هاب، مساله p - مرکز هاب، مساله بدون ظرفیت هاب و مساله پوشش هاب ارائه نمود [۳۰]. در سال ۱۹۹۷ حدادی روشی ابتکاری ساده برای حل مساله پوشش ارائه کرد [۶۲].

در همین سال روش‌های جستجوی همسایگی متعددی برای مساله p - میانه توسط هانسن^{۴۸} و ملادونویچ^{۴۹} ارائه شده است [۶۷]. ارنست^{۵۰} و کریشنامورتی^{۵۱} در سال ۱۹۹۸ با انتشار مقاله‌ای به مساله تخصیص چندوسیله‌ای p - میانه هاب بدون ظرفیت پرداختند و

^{۴۳}Titz

^{۴۴}Bart

^{۴۵}Ko

^{۴۶}Lee

^{۴۷}Karvin

^{۴۸}Hansen

^{۴۹}Mladenovic

^{۵۰}Ernst

^{۵۱}Hrishnamoorthy

الگوریتمی قطعی و همچنین الگوریتمی ابتکاری برای حل آن ارائه کردند [۴۷]. ایبری^{۵۲} در سال ۲۰۰۱ روش کارامدی برای حل مساله تخصیص تک وسیله‌ای p - هاب با سه یا چهار هاب ارائه کرد. وی در این مطالعه مثالی بزرگتر از سایر مثال‌های حل شده پیشین را حل کرد [۴۴]. فتحعلی در سال ۲۰۰۶ الگوریتم ژنتیک را برای حل مسایل p - میانه با وزن‌های مثبت و منفی ارائه کرد [۴۹]. در همین سال فتحعلی و تقی زاده الگوریتم جستجوی همسایگی را برای حل مسایل p - میانه با وزن‌های مثبت و منفی ارائه کردند [۵۰]. در سال ۲۰۰۷ گالینیر^{۵۳} و هرتز^{۵۴} مساله پوششی را مورد بررسی قرار دادند و سه الگوریتم دقیق برای حل آن ارائه نمودند. همچنین آن‌ها حالت ابتکاری این سه الگوریتم را نیز ارائه نموده و مورد آنالیز قرار دادند [۵۶]. در طی این سال‌ها هر کدام از این مسایل به‌طور گسترده و توسط بسیاری از محققین و پژوهشگران مورد مطالعه و بررسی قرار گرفتند و امروزه نتایج بسیار جالب و ارزنده‌ای در هر یک از این زمینه‌ها ارائه شده است. به‌عنوان نمونه می‌توان به مراجع [۳۵، ۱۱۰، ۱۳۲] اشاره نمود.

۲.۲.۲ سرویس‌دهندگان

فرض کنید p تعداد سرویس‌دهنده‌ها باشد. حالت‌های زیر را در نظر بگیرید:

۱. $p=1$ ؛ در این حالت فقط یک فعالیت (سرویس‌دهنده) باید مستقر شود.

۲. $p>1$ ؛ در این حالت علاوه بر پیدا کردن محل استقرار سرویس‌دهنده‌ها باید چگونگی تخصیص مشتریان به سرویس‌دهندگان نیز مشخص شوند.

در حالتی که چند سرویس‌دهنده مکان‌یابی می‌شوند، ممکن است ارتباط بین سرویس‌دهندگان جدید نیز دارای اهمیت باشد. مدل‌هایی که این هدف را برآورده می‌کنند شامل فاصله وزن‌دار بین سرویس‌دهندگان جدید هستند. به‌عنوان مثالی از این مورد می‌توان به مکان‌یابی تأسیسات نظامی اشاره کرد که در آن پراکنده شدن و فاصله داشتن سرویس‌دهندگان جدید میزان ضرر و خسارت و آسیب‌پذیری را کاهش می‌دهد. در مکان‌یابی پمپ‌های گاز و یا فست‌فودها، تصمیم‌گیرنده سعی در پراکنده کردن آن‌ها دارد تا از ایجاد رقابت مستقیم آن‌ها با یکدیگر جلوگیری کند. بر مبنای این پارامتر دو نوع مدل مختلف وجود دارد:

۱. مدل‌هایی که تعداد سرویس‌دهندگان جدید در آن ثابت است.

۲. مدل‌هایی که تعداد سرویس‌دهندگان جدید در آن متغیر است.

مدل‌های نوع دوم پیچیده‌تر از مدل‌های نوع اول هستند، زیرا باید متغیر دیگری که نشان‌دهنده تعداد سرویس‌دهندگان است به مدل اضافه شود. این مدل مکان‌یابی **شرطی**^{۵۵} نام دارد که

^{۵۲} Ebery

^{۵۳} Galinier

^{۵۴} Hertz

^{۵۵} Conditional

در آن تعدادی سرویس‌دهنده موجود هستند و هدف اضافه کردن تعدادی دیگر است. حداقل سه نوع ارتباط مختلف بین سرویس‌دهندگان می‌تواند وجود داشته باشد:

۱. فقط فاصله بین مشتریان و سرویس‌دهندگان در نظر گرفته شود، مثلاً در مکان‌یابی چند محل دفن زباله در یک کشور.
۲. فقط فاصله بین سرویس‌دهندگان محاسبه شود، مثلاً مکان‌یابی سایت‌های موشکی در یک ناحیه بدون جمعیت.
۳. هر دو نوع فاصله ذکر شده در بالا در نظر گرفته شود، مثلاً مکان‌یابی تعدادی سایت هسته‌ای در یک ناحیه دارای جمعیت.

۳.۲.۲ اندازه‌گیری فاصله

فاصله یک معیار عددی برای تشخیص میزان دور بودن اشیا است. در مسایل مکان‌یابی، فرض می‌شود که ارتباط بین تأسیسات مستقیماً وابسته به موقعیت مکانی آن‌ها باشد و معمولاً به شکل تابعی از فاصله بیان می‌شود. بنابراین مفهوم فاصله نقشی محوری در مکان‌یابی پیوسته دارد. با توجه به موقعیت و شرایط مساله اندازه‌های مختلفی برای ارزیابی فاصله به کار می‌روند. استفاده از روند نرم‌ها به عنوان تابع تخمین فاصله در یک مساله مکان‌یابی با فضای پیوسته اولین بار توسط **لاو و موریس**^{۵۶} پیشنهاد شد [۸۸]. و **ارد و وندل**^{۵۷} نیز استفاده از **نرم‌های بلوکی**^{۵۸} را مورد مطالعه قرار دادند [۱۳۵، ۱۳۶]. برای اندازه‌گیری فاصله بین مشتریان و سرویس‌دهندگان معمولاً از مترهای زیر استفاده می‌شود:

فاصله مینکوفسکی (نرم ℓ_p):^{۵۹} این فاصله بین دو نقطه $x = (x_1, x_2)$ و $y = (y_1, y_2)$ در صفحه به صورت زیر تعریف می‌شود:

$$L_p(x, y) = (k_1 |x_1 - y_1|^p + k_2 |x_2 - y_2|^p)^{1/p}; \quad 1 \leq p \leq \infty.$$

فاصله مینکوفسکی برای مدل کردن مسایلی که فاصله در جهت‌های مختلف تقارن ندارد، مناسب است. در مسایل مکان‌یابی از بین خانواده L_p ها، فاصله اقلیدسی با $p = 2$ و همچنین $k_1 = k_2 = 1$ و فاصله مستطیلی (منهتن) با $p = 1$ و $k_1 = k_2 = 1$ بیشتر مورد استفاده هستند. فاصله اقلیدسی برای مدل کردن فاصله‌های هوایی یا دریایی و فاصله مستطیلی برای مدل کردن فاصله‌های خیابانی (به شرط اینکه خیابان‌ها یک‌طرفه نباشند) مناسب هستند.

^{۵۶} Love and Morris

^{۵۷} Ward and Wendell

^{۵۸} Block norms

^{۵۹} Minkowsky

فاصله شبکه‌ای: برابر با کوتاه‌ترین مسیر بین دو نقطه روی شبکه است. برای محاسبه این فاصله الگوریتم‌های مختلفی پیشنهاد شده است که از آن جمله می‌توان الگوریتم‌های **دایکسترا^{۶۰}** و **فلوید^{۶۱}** را نام برد [۱۹].

۴.۲.۲ وزن

فاصله‌ها در تابع هدف می‌توانند دارای وزن‌های ثابت یا متغیر، یکسان یا مختلف باشند. گاهی برای مشخص کردن میزان تقاضا در هر نقطه ثابت (مشتریان) از وزن‌ها استفاده می‌شود. در مسایل شبکه‌ای، وزن یال می‌تواند نشان دهنده طول مسیر، زمان پیمایش یال، میزان مخاطره و هزینه استفاده از یال و ارائه خدمات و مواردی از این قبیل باشد.

۵.۲.۲ مشتریان

در تعدادی از مدل‌های مکان‌یابی نقش مشتری‌ها نیز منظور شده است. لذا لازم است که نحوه توزیع و میزان تقاضای آن‌ها مشخص باشد. برای مثال، ممکن است مشتریان به‌طور یکنواخت در یک ناحیه توزیع شده باشند و یا در نقاط مخصوصی از فضا، مثلاً رؤوس یک شبکه قرار داشته باشند و میزان تقاضا تابعی از فاصله باشد. مشتریان ممکن است به یک سرویس‌دهنده اختصاص یابند و یا ممکن است در انتخاب آن آزاد باشند که در آن صورت سؤال این است که آیا همیشه مشتریان به نزدیک‌ترین سرویس‌دهنده مراجعه می‌کنند یا عوامل دیگری در انتخاب سرویس‌دهنده توسط مشتری دخیل هستند.

۶.۲.۲ اهداف تصمیم‌گیرنده

مدل می‌تواند یک یا چند هدفه باشد. یعنی تصمیم‌گیرنده می‌تواند به دنبال مکان بهینه برای استقرار سرویس‌دهنده تحت یک هدف یا چند هدف به‌طور همزمان باشد. در بین مدل‌های تک هدفه حداقل سه دسته وجود دارند:

۱. کمینه‌سازی تابعی از فاصله: **کمترین مجموع^{۶۲}**، **مینیماکس^{۶۳}**.

۲. بیشینه‌سازی تابعی از فاصله: **بیشترین مجموع^{۶۴}**، **ماکزیمین^{۶۵}**.

۳. تابع‌های ترکیبی از فاصله.

^{۶۰} Dijkstra

^{۶۱} Floyd

^{۶۲} Minisum

^{۶۳} MInimax

^{۶۴} Maxisum

^{۶۵} Maximin

۳.۲ تابع هدف‌های پرکاربرد مکان‌یابی

در این بخش، لیستی از توابع هدف معروف مکان‌یابی با توضیح مختصری درباره هر کدام ارائه شده است.

ماکزین^{۶۶}: نقطه‌ای روی شبکه که مجموع فاصله‌های وزن‌دار را بیشینه کند.

مرکز: نقطه‌ای روی شبکه که بیشترین فاصله بین سرویس‌دهنده و مشتریان را کمینه کند.

مرکز کلی^{۶۷}: نقطه‌ای روی شبکه که ماکزیمم فواصل بین آن و هر نقطه‌ای روی شبکه را کمینه کند. تفاوت این مساله با مساله مرکز در این است که جواب مساله مرکز کلی لزوماً یکی از رئوس شبکه نیست و می‌تواند هر جای آن از جمله روی یال‌ها نیز قرار گیرد.

مساله چند وسیله‌ای وبر^{۶۸}: p نقطه روی صفحه که مجموع فواصل بین مشتری و نزدیک‌ترین سرویس‌دهنده را کمینه می‌کنند.

میانه^{۶۹}: نقطه‌ای روی شبکه که مجموع فواصل بین سرویس‌دهنده و مشتریان را کمینه کند.

میانه کلی^{۷۰}: نقطه‌ای روی شبکه که مجموع فواصل بین آن و هر نقطه روی شبکه را کمینه کند. تفاوت این مساله با مساله میانه در این است که جواب مساله مرکز کلی لزوماً یکی از رئوس شبکه نیست و می‌تواند هر جای آن از جمله روی یال‌ها نیز قرار گیرد.

p -میانه^{۷۱}: استقرار p نقطه روی شبکه به‌گونه‌ای که مجموع فواصل بین مشتری تا نزدیک‌ترین سرویس‌دهنده کمینه گردد.

p -مرکز^{۷۲}: استقرار p نقطه روی شبکه به‌گونه‌ای که بیشترین فاصله بین مشتری و سرویس‌دهنده‌ها کمینه گردد.

۴.۲ برخی مسایل مهم مکان‌یابی

در این بخش برخی مدل‌های مسایل مکان‌یابی پرکاربرد آمده است.

^{۶۶}Maxian

^{۶۷}General Center

^{۶۸}Multi facility Weber Problem

^{۶۹}Median

^{۷۰}General Median

^{۷۱}P-median

^{۷۲}P-center

۱.۴.۲ مساله مکان‌یابی چند وسیله‌ای پیوسته

به‌طور کلی، مسایل مکان‌یابی پیوسته (دو بعدی) به دنبال یافتن مکان یک یا چند سرویس دهنده بر روی صفحه هستند. شکل کلی این مساله به صورت زیر است [۸۸]:

$$\min W(X) = \sum_{i=1}^m \sum_{j=1}^n w_{1ij} d(X_i, A_j) + \sum_{i=1}^{m-1} \sum_{r=i+1}^m w_{2ir} d(X_i, X_r). \quad (1.2)$$

که در آن $X = (X_1, X_2, \dots, X_m)$ مختصات مکان سرویس-دهنده‌ها و $j = 1, \dots, n, A_j = (a_{j1}, a_{j2})$ مراکز تقاضا هستند. $i = 1, \dots, m, X_i = (x_{i1}, x_{i2})$ و $i = 1, \dots, m, j = 1, \dots, n, w_{1ij}$ و $i = 1, \dots, m, r = i+1, \dots, m, w_{2ir}$ به ترتیب هزینه انتقال کالا از سرویس دهنده i - ام به مرکز تقاضای j - ام و هزینه انتقال کالا از سرویس دهنده i - ام به سرویس دهنده r - ام هستند. همچنین $d(X_i, A_j)$ و $d(X_i, X_r)$ به ترتیب فاصله بین سرویس دهنده i - ام و مرکز تقاضای j - ام و فاصله بین دو مرکز سرویس دهی i - ام و r - ام هستند.

اگر فرض شود که $m = 1$ ، در این صورت مساله ۱.۲ به مساله تک وسیله‌ای زیر تبدیل می‌شود:

$$\min W(X) = \sum_{j=1}^n w_j d(X, A_j). \quad (2.2)$$

مساله اخیر یکی از شناخته شده‌ترین مسایل مکان‌یابی پیوسته است که به مساله تک وسیله‌ای فرما - وبر معروف است. در این مساله n نقطه در صفحه موجود است. هدف یافتن نقطه‌ای جدید در صفحه است به‌طوری‌که مجموع کل فواصل نقاط موجود تا نقطه بهینه کمینه شود [۸۸]. همچنین اگر هدف پیدا کردن X به گونه‌ای باشد که فاصله وزنی X تا دورترین نقطه موجود کمینه شود، آن‌گاه مساله را مساله تک وسیله‌ای مینیماکس گویند که به صورت زیر است:

$$\min \max_{i=1, \dots, n} w_i d(X, A_i).$$

در سال ۱۹۳۷ **وایزفیلد**^{۷۳} روشی تکراری برای حل مساله فرما - وبر ارائه کرد [۱۳۸]. میهل در سال ۱۹۵۸ برای اولین بار الگوریتم تکراری وایزفیلد را برای مساله مکان‌یابی چند وسیله‌ای با نرم اقلیدسی گسترش داد [۸۸]. در سال ۱۹۶۴ **فرانسیس**^{۷۴} مساله چند وسیله‌ای با نرم مستطیلی را مورد بررسی قرار داد [۵۴]. **وسولوزکی**^{۷۵} و **لاو**^{۷۶} در سال ۱۹۷۱ به مکان‌یابی در صفحه با فاصله مستطیلی در بین مقاصد نقاط و نواحی مستطیلی پرداختند [۱۴۰]. همچنین **وسولوزکی** در سال ۱۹۷۴ مساله مکان‌یابی خط بهینه در صفحه را بیان نمود [۱۳۹]. در

^{۷۳}Weiszfeld

^{۷۴}Francis

^{۷۵}Wesolowsky

^{۷۶}Love

سال ۱۹۷۸ درزنر^{۷۷} و وسولوزکی مساله مکان‌یابی روی سطح کروی را بررسی کردند [۴۳]. همچنین در سال‌های اخیر نیز برخی از پژوهشگران و محققان مطالعه در این زمینه را ادامه دادند که منجر به حصول نتایج جدیدی شده است [۲۲، ۲۹].

در حالت گسسته مکان و سائل جدید تنها می‌تواند در نقاط موجود قرار گیرند و حالت خاصی از آن وقتی است که نقاط موجود روی یک شبکه قرار داشته و مکان و سائل جدید نیز باید نقطه‌ای از شبکه باشند. در این حالت وقتی تابع هدف به‌صورت کمترین مجموع و یا مینیماکس است، به ترتیب مساله را مساله میانه و مرکز می‌نامند. همچنین اگر هدف پیدا کردن مکان p وسیله جدید باشد، آنگاه این مسایل به ترتیب مسایل p - میانه و p - مرکز نامیده می‌شوند.

۲.۴.۲ مساله p - میانه و مدل ریاضی آن

در مساله p - میانه روی شبکه‌ها، هدف پیدا کردن یک مجموعه شامل p - سرویس‌دهنده $X = \{x_1, x_2, \dots, x_p\}$ روی شبکه $N = (V, E)$ با مجموعه راس‌های V و مجموعه یال‌های E است به طوری که اگر هر راس v_i دارای وزن w_i باشد، مجموع فواصل وزنی این مجموعه تا راس‌های روی شبکه N مینیمم شود، یعنی:

$$\min_{X \in N} F(X) = \sum_{v_i \in V} w_i d(X, v_i)$$

که در آن $d(X, v_i) = \min_{x_j \in V} d(x_j, v_i)$ و $d(v_i, v_j)$ اندازه کوتاه‌ترین مسیر بین دو راس v_i و v_j است.

در حالتی که $x \in X$ بتواند هر نقطه‌ای از شبکه (روی راس یا یال) باشد، مساله فوق p - میانه مطلق نامیده می‌شود، اما اگر هدف پیدا کردن نقاط میانه فقط روی راس‌ها مورد نظر باشد، مساله را p - میانه راسی نامند.

قضیه ۱.۴.۲. [۶۳] حداقل یک مجموعه متشکل از p رأس در مجموعه جواب‌های مساله p - میانه (با وزن راس‌های مثبت) وجود دارد.

کریو^{۷۸} و حکیمی نشان دادند که مساله p - میانه یک مساله NP - سخت است [۷۸] و لذا برای p های بزرگ، پیچیدگی زمانی این مساله زیاد می‌باشد. همچنین، روش‌های مختلفی برای حل مسایل p - میانه ارائه شده است [۷۸، ۱۲۶، ۸۹].

۳.۴.۲ مساله مکان‌یابی p - مرکز

مساله مکان‌یابی p - مرکز به تعیین مکان p سرویس‌دهنده (مرکزها) $X = \{x_1, x_2, \dots, x_p\}$ روی شبکه $N = (V, E)$ می‌پردازد به طوری که ماکزیمم فاصله هر مشتری (نقاط تقاضا) تا

^{۷۷}Drezner

^{۷۸}Kariv

نزدیک‌ترین سرویس‌دهنده مینیمم شود. تابع هدف این مساله به صورت زیر است:

$$\min_{X \in N} G(X) = \max_{v_i \in V} \{w_i d(X, v_i)\}$$

در حالتی که $x \in X$ بتواند هر نقطه‌ای از شبکه (روی راس یا یال) باشد، مساله فوق p - مرکز مطلق نامیده می‌شود، اما اگر هدف پیدا کردن نقاط مرکز فقط روی راس‌های شبکه مورد نظر باشد، مساله را p - مرکز راسی نامند. مساله p - مرکز نیز یک مساله NP - سخت است [۵۷] و الگوریتم‌های مختلفی برای حل آن نیز ارائه شده است [۷۸، ۱۰۵].

فصل ۳

مسایل مکان‌یابی تک‌وسایله‌ای توافقی تحت تابع زیان متقارن

۱.۳ مقدمه

در طول دهه‌های اخیر، توجه و علاقه‌مندی به مسایل محیطی و دنیای واقعی افزایش پیدا کرده است. این مسایل عمدتاً در نتیجه پیشرفت‌های صنعتی و فناوری ایجاد شده‌اند و غالباً تأثیرات منفی در محیط و مشکلات زیست‌محیطی را به دنبال می‌آورند.

مسایلی مانند استقرار یک سرویس‌دهنده که با وجود تأمین خدمات و تقاضای مشتریان، اثرات نامطلوبی مثل سروصدا، گردوخاک، گازهای سمی، بوی بد، شلوغی، تشعشعات هسته‌ای و غیره را بر مشتریان تحمیل می‌کند، از این دسته هستند. تصمیم‌گیری در مورد احداث چنین سرویس‌دهندگانی در دنیای واقعی و زندگی، کار پیچیده و دشواری بوده و تعداد آن‌ها نیز قابل توجه است. مکان‌یابی چنین سرویس‌دهندگانی که هم دارای تأثیرات مثبت و هم تأثیرات منفی بر روی مشتریان هستند، مکان‌یابی توافقی نام‌گذاری شده است.

بنابراین با توجه به اهمیت موضوع و وجود بسیاری از مسایل در دنیای واقعی که با استفاده از مفهوم مکان‌یابی توافقی به‌صورت ریاضی مدل می‌شوند، در این فصل به بررسی یک مساله مکان‌یابی تک‌وسایله‌ای توافقی پرداخته می‌شود.

این فصل در پنج بخش تهیه شده است. در بخش دوم، مبانی نظری و پیشینه تحقیق بیان

می‌شود. در بخش سوم، مساله مکان‌یابی تک‌وسيله‌ای توافقی با در نظر گرفتن توابع خطای مختلف، به‌صورت ریاضی مدل‌سازی می‌شود. در بخش چهارم، قضایایی برای وجود جواب در مدل‌های پیشنهادی، اثبات و برای حل آن‌ها الگوریتم‌های شبه‌وایزفیلد (WLA)^۱، گاوس-نیوتن (GN)^۲ و فراابتکاری رقابت استعماری (ICA)^۳ بیان می‌شود. در بخش پایانی نیز مدل‌های ارائه‌شده در بخش سوم و چهارم روی چندین مثال عددی بررسی و تحلیل می‌شوند.

۲.۳ مبانی نظری و پیشینه تحقیق

مسایل مکان‌یابی در توابع هدف، فواصل مورد استفاده، تعداد و اندازه سرویس‌دهنده‌هایی که باید تأسیس شوند و همچنین در چندین عامل دیگر متفاوت هستند. هونگ ژونگ جیا و همکاران هشت عامل را که در دسته‌بندی مدل‌های مکان‌یابی تسهیلات مؤثر هستند، معرفی کردند [۷۲]. این هشت عامل عبارت‌اند از: مشخصات جغرافیایی، خصوصیات تسهیلات، اهداف، روش حل، الگوهای تقاضا، انواع زنجیره عرضه، افق زمانی و مولفه‌های ورودی. بنابراین، بسته به نوع هدف و با استفاده از هر یک از این شاخص‌ها، مدل‌های گوناگون مساله مکان‌یابی به دست می‌آید. از این‌رو، مسایل مکان‌یابی پیوسته نوع خاصی از مسایل مکان‌یابی هستند که در آن‌ها هدف یافتن مکان یک یا چند سرویس‌دهنده روی صفحه است. یکی از شناخته‌شده‌ترین و مهم‌ترین مسایل مکان‌یابی پیوسته، مساله مکان‌یابی تک‌وسيله‌ای فرما-وبر است. در این مساله نقاطی در صفحه موجود بوده و هدف یافتن نقطه‌ای جدید در صفحه است؛ به‌طوری‌که مجموع کل فواصل نقاط موجود تا نقطه جدید کمینه شود [۸۸]. مدل ریاضی یک مساله فرما-وبر را می‌توان به‌صورت رابطه (۱.۳) بیان کرد.

$$\text{Min} \sum_{i=1}^n w_i d(X, A_i). \quad (1.3)$$

در رابطه (۱.۳) منظور از $d(X, A_i)$ فاصله بین سرویس‌دهنده X و نقاط تقاضای A_i است. w_i نیز وزن مربوط به هر نقطه تقاضای A_i است. نمونه‌های گوناگون از مساله مکان‌یابی وبر را می‌توان در مراجع [۲۴]، [۳۳] و [۱۳۰] مشاهده نمود. وایزفیلد در سال ۱۹۳۷ یک الگوریتم گرادانی برای حل مساله فرما-وبر ارائه کرد [۱۳۸].

در ادامه، بسیاری از محققان از الگوریتم وایزفیلد برای حل مدل‌های گوناگون مسایل مکان‌یابی استفاده نمودند. به‌طور مثال، میهل^۴ در سال ۱۹۵۸ الگوریتم وایزفیلد را برای مسایل مکان‌یابی چندوسيله‌ای و با نرم اقلیدسی گسترش داد [۹۲]. ایستر^۵ و همکاران در سال ۱۹۷۳ الگوریتم وایزفیلد را با استفاده از روش تقریبی هیپربولوئید برای مسایل مکان‌یابی

^۱Weiszfeld-Like Algorithm

^۲Gauss Newton Algorithm

^۳Imperialist Competitive Algorithm

^۴Miehle

^۵Eyster

چندوسیله‌ای با نرم‌های خطی و اقلیدسی گسترش دادند [۴۸]. موریس و وردینی^۶ نیز در سال ۱۹۷۳ و موریس در سال ۱۹۸۱ به ترتیب در مقالاتی مجزا، الگوریتم وایزفیلد را برای مسایل مکان‌یابی با نرم ℓ_p ، مورد بحث قرار دادند و الگوریتم‌هایی شبه‌وایزفیلد برای مسایل خود ارائه کردند [۹۹] [۱۰۰]. واردی^۷ و ژانگ^۸ در سال ۲۰۰۱ الگوریتم وایزفیلد را برای حل مساله فرما-وبر اصلاح کردند [۱۳۳]. بعدها در سال ۲۰۰۹ بریمبرگ^۹ و همکاران در مقاله‌ای الگوریتم وایزفیلد را برای یک مساله تک‌وسیله‌ای با تابع هدف کمترین مجموع و تحت نرم‌های ترکیبی ℓ_p گسترش دادند [۲۵]. همچنین لایجن^{۱۰} و بن اسرائیل^{۱۱} در سال ۲۰۱۰ الگوریتم وایزفیلد را برای مسایل تخصیص مورد استفاده قرار دادند [۷۴]. فتحعلی نیز در سال ۲۰۱۴ مساله مکان‌یابی پشتیبان چندوسیله‌ای در صفحه را با ارائه یک الگوریتم شبه‌وایزفیلد حل کرد [۵۱].

از طرف دیگر باید توجه داشت که در طول دهه‌های اخیر تلاش‌های بسیار زیادی برای ایجاد مدل‌های مکان‌یابی که مشخصه‌های بیشتری از دنیای واقعی را در نظر می‌گیرند، انجام گرفته است. یکی از کاربردی‌ترین مسایلی که در دنیای واقعی و در زندگی شهری وجود دارد، مکان‌یابی سرویس‌دهندگانی است که دارای تأثیرات مثبت و تأثیرات منفی بر روی مشتریان هستند. فتحعلی و همکاران یک مساله مکان‌یابی وبر را با در نظر گرفتن این خصوصیت که یک سرویس‌دهنده می‌تواند هم‌زمان دارای تأثیرات مطلوب و نامطلوب باشد، مورد بررسی قرار دادند و این مساله را مساله مکان‌یابی توافقی نامیدند [۵۳]. آن‌ها در این مساله برای هر نقطه تقاضا، یک شعاع توافقی را در نظر گرفته و فاصله بین تسهیلات جدید و نقاط تقاضا را برابر با شعاع توافقی متناظر با هر نقطه تقاضا تعیین کردند. آن‌ها با استفاده از روش هندسی مربع بزرگ مربع کوچک^{۱۲} الگوریتمی برای حل مساله تحت نرم اقلیدسی پیشنهاد دادند. سپس جمالیان و فتحعلی یک مدل برنامه‌ریزی خطی برای مساله با هدف کمینه‌کردن مجموع وزنی قدرمطلق خطا، ارائه کردند. اخیراً فتحعلی و جمالیان به بررسی مساله کمینه‌کردن مجموع مربعات خطا پرداخته و آن را مساله مکان‌یابی وبر مربعی توافقی (GSWLP)^{۱۳} نامیدند [۵۲]. آن‌ها با استفاده از الگوریتم شبیه‌سازی پرندگان به حل مساله فوق تحت نرم اقلیدسی پرداختند.

^۶Verdini

^۷Vardi

^۸Zhang

^۹Brimberg

^{۱۰}Iyigun

^{۱۱}Ben-Israel

^{۱۲}big Square Small square

^{۱۳}Goal Square Weber Location Problem

۳.۳ مساله مکان‌یابی تک‌وسیله‌ای توافقی با نرم ℓ_p

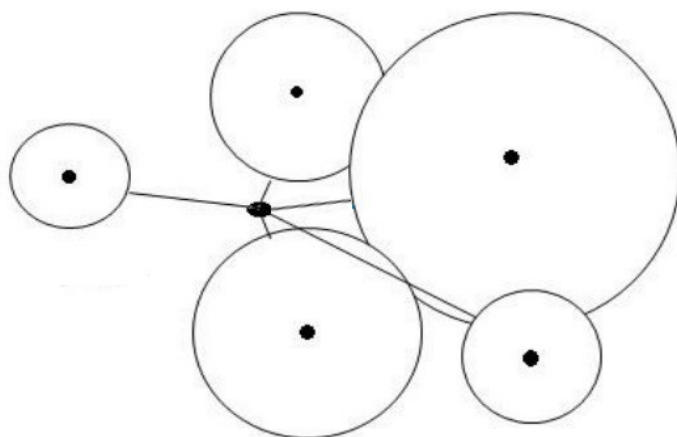
۱.۳.۳ بیان مساله

در یک مساله مکان‌یابی توافقی، برای هر مشتری یک شعاع توافقی در نظر گرفته شده است و هدف پیدا کردن مکانی برای سرویس‌دهندگان است؛ به‌طوری‌که فاصله آن‌ها تا مشتریان دقیقاً برابر با شعاع توافقی متناظر با هر مشتری باشد. اما از آنجایی‌که در واقعیت ممکن است چنین نقطه‌ای پیدا نشود، این مکان (نقطه) چنان برآورد می‌شود که فاصله آن تا نقاط تقاضا، تا حد امکان به شعاع‌های توافقی متناظر نزدیک شود. انتخاب این مکان برآورد شده، ضرر و زیان‌هایی (خطایی) را نسبت به نقطه ایده‌آل در پی خواهد داشت. هدف در این مساله، کمینه‌شدن زیان ناشی از انتخاب این مکان در صفحه است.

به‌عنوان مثالی از این موضوع، فرض کنید که باید فروشگاه‌هایی در یک شهر احداث شود. با توجه به ساختار زندگی شهری، نحوه و توزیع جمعیت، هر چه فروشگاه به مرکز شهر و مناطق پرجمعیت نزدیکتر باشد، از یک طرف دسترسی ساکنان افزایش یافته و متوسط هزینه حمل‌ونقل آنها کاسته شده و تقاضای مشتریان به خوبی برآورده می‌شود و از طرف دیگر، احداث فروشگاه در چنین جایگاه‌هایی هزینه‌های زیادی (خرید زمین و مالیات و...) در پی خواهد داشت. تصمیم‌گیری برای احداث چنین سرویس‌دهندگانی در دنیای واقعی کار پیچیده و دشواری است. برای مدل‌سازی این مسایل، یک کران مشترک (شعاع توافقی) برای دوری و نزدیکی از مرکز شهر در نظر گرفته می‌شود که این کران مشترک، معیاری برای جلب رضایت مشتریان و مدیران است؛ لذا با تعیین یک کران مشخص برای فاصله از مرکز شهر، مساله با هدف کمترین مجموع خطای فاصله از کران‌های مذکور مدل می‌شود. در حقیقت، میزان خطای دوری از این کران کمینه می‌شود. نمایش گرافیکی مساله برای پنج نقطه تقاضا و برای شعاع‌های مختلف در شکل ۱.۳ نشان داده شده است.

۲.۳.۳ مدل‌سازی مساله

فرض کنید n نقطه تقاضای (مشتریان) $A_1, \dots, A_n$ در صفحه موجود باشد که متناسب با هر نقطه تقاضای A_i ، یک شعاع توافقی R_i و یک وزن مثبت w_i وجود دارد. در این صورت با توجه به تعریف مساله مکان‌یابی تک‌وسیله‌ای توافقی، هدف یافتن یک مکان جدید است؛ به‌طوری‌که فاصله وزنی بین این نقطه جدید و نقاط تقاضای A_i ، تا حد امکان برابر با شعاع توافقی R_i شود. در حقیقت باید میزان خطای دوری از این کران کمینه شود. فرض کنید که $A_i = (a_i, b_i)$ مختصات نقاط تقاضا و $X = (x, y)$ مختصات مکانی مجهولی باشد که باید سرویس‌دهنده موردنظر مکان‌یابی شود، در این صورت مساله مکان‌یابی تک‌وسیله‌ای توافقی



شکل ۱.۳: نمای گرافیکی مساله مکان‌یابی توافقی

با نرم ℓ_p به صورت رابطه (۲.۳) مدل‌سازی می‌شود:

$$\min f(X) = \sum_{i=1}^n w_i \cdot e(d(X, A_i) - R_i) \quad (2.3)$$

که w_i وزن مربوط به هر نقطه تقاضای A_i ، R_i شعاع توافقی متناظر با نقطه تقاضای A_i است و $e(d(X, A_i) - R_i)$ مفهوم خطا را بیان می‌کند که $d(X, A_i)$ فاصله بین دو نقطه تحت نرم ℓ_p است؛ یعنی:

$$d(X, A_i) = (|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}}. \quad (3.3)$$

حال اگر تابع خطا در رابطه (۲.۳)، به صورت تابع مجموع مربعات خطا در نظر گرفته شود، مساله مکان‌یابی تک‌وسيله‌ای توافقی با مجموع مربعات خطا به شکل (۴.۳) بدست می‌آید:

$$\min f(X) = \sum_{i=1}^n w_i \cdot (d(X, A_i) - R_i)^2 \quad (4.3)$$

و اگر تابع خطای $e(d(X, A_i) - R_i)$ ، یک تابع خطای قدرمطلق در نظر گرفته شود، در این صورت، مساله با مجموع خطای قدر مطلق به شکل (۵.۳) مدل‌سازی می‌شود:

$$f(X) = \sum_{i=1}^n w_i \cdot |d(X, A_i) - R_i| \quad (5.3)$$

مسائل مکان‌یابی تک‌وسيله‌ای توافقی با توابع هدف (۴.۳) و (۵.۳) دارای جواب بهینه هستند و قضایای ۱.۳.۳ و ۲.۳.۳ برای آن‌ها برقرار است.

تعریف ۱.۳.۳. قرار دهید

$$\begin{cases} a_{\min} = \min\{a_i - R_i | i = 1, \dots, n\} \\ a_{\max} = \max\{a_i + R_i | i = 1, \dots, n\} \end{cases}, \begin{cases} b_{\min} = \min\{b_i - R_i | i = 1, \dots, n\} \\ b_{\max} = \max\{b_i + R_i | i = 1, \dots, n\} \end{cases}$$

همچنین مجموعه نقاط $RH_3 = (a_{\max}, b_{\max})$ ، $RH_2 = (a_{\min}, b_{\max})$ ، $RH_1 = (a_{\min}, b_{\min})$ و $RH_4 = (a_{\max}, b_{\min})$ را در نظر بگیرید. پوسته گسترش‌یافته مستطیلی مجموعه نقاط فوق به صورت کوچکترین مستطیل شامل این نقاط با اضلاع موازی محورهای مختصات تعریف می‌شود.

قضیه ۱.۳.۳. جواب بهینه مساله مکان‌یابی تک‌وسيله‌ای توافقی (۴.۳) در پوسته گسترش‌یافته مستطیلی نقاط قرار دارد.

برهان. فرض کنید $X = (x, y)$ خارج از پوسته مستطیلی نقاط RH_1 ، RH_2 ، RH_3 و RH_4 قرار داشته باشد. در حالت اول، فرض کنید $x > a_{\max}$ ؛ اگر $X' = (a_{\max}, y)$ نتیجه می‌شود:

$$d(X, A_i) = (|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} > (|a_{\max} - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} = d(X', A_i) > R_i$$

بنابراین،

$$\left((|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} - R_i \right)^2 > \left((|a_{\max} - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} - R_i \right)^2$$

و با توجه به مثبت بودن ضرایب w_i ، به دست می‌آید که $f(X) > f(X')$ و لذا $X = (x, y)$ نمی‌تواند جواب بهینه باشد. برای حالت‌هایی که $x < a_{\min}$ ، $y < b_{\min}$ و $y > b_{\max}$ نیز به صورت کاملاً مشابه اثبات می‌شود که نقاط X نمی‌توانند جواب‌های بهینه باشند. بنابراین، جواب بهینه مساله در پوسته مستطیلی حاصل از نقاط RH_1 ، RH_2 ، RH_3 و RH_4 قرار دارد. $\square$

قضیه ۲.۳.۳. جواب بهینه مساله مکان‌یابی تک‌وسيله‌ای توافقی (۵.۳) در پوسته گسترش‌یافته مستطیلی نقاط $A_i = (a_i, b_i)$ قرار دارد.

برهان. فرض کنید $X = (x, y)$ خارج از پوسته مستطیلی نقاط RH_1 ، RH_2 ، RH_3 و RH_4 قرار داشته باشد. در حالت اول، فرض کنید $y > b_{\max}$ ، اگر $X' = (x, b_{\max})$ آن‌گاه:

$$d(X, A_i) = (|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} > (|x - a_i|^p + |b_{\max} - b_i|^p)^{\frac{1}{p}} = d(X', A_i) > R_i$$

بنابراین،

$$\left| (|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} - R_i \right| > \left| (|x - a_i|^p + |b_{\max} - b_i|^p)^{\frac{1}{p}} - R_i \right|$$

و با توجه به مثبت بودن ضرایب w_i به دست می‌آید:

$$\sum_{i=1}^n \left(w_i \left| (|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} - R_i \right| \right) > \sum_{i=1}^n \left(w_i \left| (|x - a_i|^p + |b_{\max} - b_i|^p)^{\frac{1}{p}} - R_i \right| \right)$$

پس نتیجه گرفته می‌شود که $f(X) > f(X')$ و بنابراین، $X = (x, y)$ نمی‌تواند جواب بهینه باشد. برای حالت‌هایی که $x < a_{\min}$ ، $x > a_{\max}$ و $y < b_{\min}$ نیز به صورت کاملاً مشابه اثبات می‌شود که نقاط X نمی‌توانند جواب‌های بهینه باشند. بنابراین، نتیجه می‌گیریم که جواب بهینه مساله در پوسته مستطیلی حاصل از نقاط RH_1 ، RH_2 ، RH_3 و RH_4 قرار دارد. $\square$

لم ۱.۳.۳. توابع هدف مسایل مکان‌یابی (۴.۳) و (۵.۳) نامحدب هستند.

□

برهان. [۵۳]

ملاحظه ۱.۳.۳. مشتقات توابع هدف مسایل مکان‌یابی (۴.۳) و (۵.۳) ممکن است در هر جایی از صفحه تعریف شده نباشند.

۴.۳ الگوریتم‌های حل مساله مکان‌یابی تک وسیله‌ای توافقی

در این بخش، مساله مکان‌یابی تک وسیله‌ای (۴.۳) با استفاده از سه الگوریتم شبه‌وایزفیلد، گاوس-نیوتن و فراابتکاری رقابت استعماری حل می‌شود. مساله مکان‌یابی (۵.۳) نیز مشابه مساله (۴.۳) حل خواهد شد.

۱.۴.۳ الگوریتم شبه‌وایزفیلد

با فرض این که جواب بهینه (۴.۳) در نقطه مشتق پذیر $X^* = (X_1^*, X_2^*)$ اتفاق می‌افتد از شرط لازم بهینگی داریم:

$$\frac{\partial F(X^*)}{\partial X_k} = 2 \sum_{i=1}^n \frac{w_i (d(X^*, A_i) - R_i) \operatorname{sign}(X_k^* - a_{ik}) |X_k^* - a_{ik}|^{p-1}}{d(X^*, A_i)^{p-1}} = 0, \quad k = 1, 2$$

که در آن $A_i = (a_{i1}, a_{i2})$. با توجه به این که $\operatorname{sign}(X_k^* - a_{ik}) |X_k^* - a_{ik}| = X_k^* - a_{ik}$ بدست می‌آوریم:

$$X_k^* = \frac{\sum_{i=1}^n \frac{w_i (d(X^*, A_i) - R_i) |X_k^* - a_{ik}|^{p-2} a_{ik}}{d(X^*, A_i)^{p-1}}}{\sum_{i=1}^n \frac{w_i (d(X^*, A_i) - R_i) |X_k^* - a_{ik}|^{p-2}}{d(X^*, A_i)^{p-1}}} \quad k = 1, 2 \quad (6.3)$$

بنابراین با استفاده از تکنیک نقطه ثابت و جای‌گذاری $X^{(t)}$ و $X^{(t+1)}$ در طرفین معادله (۶.۳) دنباله تکراری زیر را بدست می‌آوریم:

$$X_k^{(t+1)} = \frac{\sum_{i=1}^n \frac{w_i (d(X^{(t)}, A_i) - R_i) |X_k^{(t)} - a_{ik}|^{p-2} a_{ik}}{d(X^{(t)}, A_i)^{p-1}}}{\sum_{i=1}^n \frac{w_i (d(X^{(t)}, A_i) - R_i) |X_k^{(t)} - a_{ik}|^{p-2}}{d(X^{(t)}, A_i)^{p-1}}} \quad k = 1, 2 \quad (7.3)$$

دنباله تکراری (۷.۳) وقتی $X_k^{(t)} = a_{ik}$ در حالت $p \leq 2$ و همچنین برای $X^{(t)} = A_i$ در حالت $p > 2$ تعریف نشده است. برای رفع این مشکل تقریب هموار $f^h(X)$ را به جای $f(X)$ در نظر

می‌گیریم، این ایده برای اولین بار در [۸۸] مورد استفاده قرار گرفته است:

$$\min f^h(X) = \sum_{i=1}^n w_i (d'(X, A_i) - R_i)^2 \quad (۸.۳)$$

که در آن، $d'(X, A_i) = \left(\left((X_1 - a_{i1})^2 + \varepsilon \right)^{\frac{p}{2}} + \left((X_2 - a_{i2})^2 + \varepsilon \right)^{\frac{p}{2}} \right)^{\frac{1}{p}}$ و ε عدد مثبت کوچکی است.

لم ۱.۴.۳. تابع هدف تقریب زده شده مساله تک‌وسيله‌ای توافقی با نرم ℓ_p و با مجموع مربعات خطا یعنی $f^h(X)$ ، به تابع هدف مساله اصلی یعنی $f(X)$ همگرا است و همچنین

$$|f^h(X) - f(X)| \leq \frac{1}{p} \varepsilon \left(\sum_{i=1}^n w_i \right) + \frac{p+1}{p} \varepsilon^{\frac{1}{p}} \left(\sum_{i=1}^n (w_i R_i) \right) \quad (۹.۳)$$

برهان. از (۴.۳) و (۸.۳) به دست می‌آید:

$$|f^h(X) - f(X)| = \left| \sum_{i=1}^n w_i \left(\left((x - a_i)^2 + \varepsilon \right)^{\frac{p}{2}} + \left((y - b_i)^2 + \varepsilon \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} - R_i \right| - \sum_{i=1}^n w_i \left((|x - a_i|^p + |y - b_i|^p)^{\frac{1}{p}} - R_i \right)^2 \right|$$

با قرار دادن $s_1 = x - a_i$ ، $s_2 = y - b_i$ ، $s_3 = \varepsilon^{\frac{1}{p}}$ و با استفاده از نامساوی مینکوفسکی نتیجه می‌شود که $|f^h(X) - f(X)| \leq \frac{1}{p} \varepsilon \left(\sum_{i=1}^n w_i \right) + \frac{p+1}{p} \varepsilon^{\frac{1}{p}} \left(\sum_{i=1}^n (w_i R_i) \right)$. بنابراین وقتی $\varepsilon \rightarrow 0$ ، نتیجه می‌شود که $\max \{|f^h(X) - f(X)|\} \rightarrow 0$. $\square$

از شرط لازم بهینگی برای مساله (۸.۳) داریم:

$$\frac{\partial F(X^*)}{\partial X_k} = \frac{2}{p} \sum_{i=1}^n \frac{w_i (d'(X^*, A_i) - R_i) (X_k^* - a_{ik})}{\left((X_k^* - a_{ik})^2 + \varepsilon \right)^{1 - \frac{p}{2}} d'(X^*, A_i)^{p-1}} = 0 \quad k = 1, 2 \quad (۱۰.۳)$$

با فرض $d''(X, A_i) = d'(X, A_i)^{p-1}$ و $d'''(X_k, a_{ik}) = \left((X_k - a_{ik})^2 + \varepsilon \right)^{1 - \frac{p}{2}}$ داریم:

$$X_k^* = \frac{\sum_{i=1}^n \frac{w_i (d'(X^*, A_i) - R_i) a_{ik}}{d'''(X_k^*, a_{ik}) d''(X^*, A_i)}}{\sum_{i=1}^n \frac{w_i (d'(X^*, A_i) - R_i)}{d'''(X_k^*, a_{ik}) d''(X^*, A_i)}} \quad k = 1, 2 \quad (۱۱.۳)$$

با استفاده از تکنیک نقطه ثابت دنباله تکراری زیر بدست می‌آید:

$$X_k^{(t+1)} = \frac{\sum_{i=1}^n \frac{w_i (d'(X^{(t)}, A_i) - R_i) a_{ik}}{d'''(X_k^{(t)}, a_{ik}) d''(X^{(t)}, A_i)}}{\sum_{i=1}^n \frac{w_i (d'(X^{(t)}, A_i) - R_i)}{d'''(X_k^{(t)}, a_{ik}) d''(X^{(t)}, A_i)}} \quad k = 1, 2 \quad (۱۲.۳)$$

واضح است که دنباله تکراری (۱۲.۳) در $X_k^{(t)} = a_{ik}$ و $X^{(t)} = A_i$ تعریف شده است. رابطه (۱۲.۳) را می‌توان به صورت زیر نوشت:

$$(X_1^{(t+1)}, X_p^{(t+1)}) = (X_1^{(t)}, X_p^{(t)}) - \frac{1}{W'} \nabla f(X_1^{(t)}, X_p^{(t)}) \quad (۱۳.۳)$$

که در آن

$$W' = 2 \sum_{i=1}^n \frac{w_i (d'(X^{(t)}, A_i) - R_i)}{d'''(X^{(t)}, a_{ik}) d''(X^{(t)}, A_i)}.$$

ملاحظه ۱.۴.۳. اگر $W' > 0$ آن‌گاه $(X_1^{(t+1)}, X_p^{(t+1)})$ روی خط تندترین کاهش قرار دارد و اگر $W' < 0$ آن‌گاه $(X_1^{(t+1)}, X_p^{(t+1)})$ روی خط تندترین افزایش قرار دارد. بنابراین اگر $W' > 0$ از $(X_1^{(t+1)}, X_p^{(t+1)})$ و اگر $W' < 0$ از $(X_1^{(t)}, X_p^{(t)}) - (X_1^{(t+1)}, X_p^{(t+1)})$ به عنوان نقطه شروع تکرار بعد استفاده می‌کنیم. وقتی که $W' = 0$ تکرار بعدی تعریف نشده است، البته احتمال این که $W' = 0$ بسیار کم می‌باشد.

سه لم زیر برای (۱۲.۳) بیان و اثبات می‌شود.

لم ۲.۴.۳. اگر $X^{(t+1)} = X^{(t)}$ باشد، آن‌گاه $\frac{\partial f^h(X^{(t)})}{\partial X} = 0$ است.

برهان. با توجه به رابطه (۱۳.۳) و ملاحظه ۱.۴.۳، اگر $X^{(t+1)} = X^{(t)}$ باشد، در این صورت، $\frac{\partial f^h(X^{(t)})}{\partial X} = 0$ و در نتیجه، $d^{(t)T} \cdot \frac{\partial f^h(X^{(t)})}{\partial X} = 0$ است. $\square$

لم ۳.۴.۳. اگر $X^{(t+1)} \neq X^{(t)}$ باشد، آن‌گاه $f^h(X^{(t+1)}) < f^h(X^{(t)})$ است.

برهان. با توجه به رابطه (۱۳.۳) و ملاحظه ۱.۴.۳، اگر $X^{(t+1)} \neq X^{(t)}$ در این صورت، $\frac{\partial f^h(X^{(t)})}{\partial X} \neq 0$ است. با قرار دادن $d = -\frac{\partial f^h(X^{(t)})}{\partial X}$ نتیجه گرفته می‌شود که $d^T \cdot \nabla f^h(X^{(t)}) < 0$ و لذا $f^h(X^{(t+1)}) < f^h(X^{(t)})$. $\square$

لم ۴.۴.۳. اگر $f^h(X^{(t+1)}) = f^h(X^{(t)})$ آن‌گاه $X^{(t+1)} = X^{(t)}$

برهان. $f^h(X^{(t+1)}) = f^h(X^{(t)})$ بدین معنی است که برای هر t ، مسیر $d^{(t)}$ یک مسیر غیرکاهشی است و از لم ۳.۴.۳ نتیجه گرفته می‌شود که $X^{(t+1)} = X^{(t)}$. $\square$

با توجه به مطالب گفته‌شده الگوریتم شبه‌وایزفیلد به صورت زیر بیان می‌شود.

الگوریتم ۴ شبه‌وایزفیلد (WLA)

- ۱: $X^{(0)}$ ، یک تقریب اولیه از جواب بهینه انتخاب کن و قرار بده $k = 0$.
- ۲: تا زمان برآورده نشدن شرط توقف، کارهای زیر را انجام بده:
- ۳: $X^{(t+1)}$ را از رابطه (۱۲.۳) محاسبه کن.
- ۴: اگر $|f_h^{(k)} - f_h^{(k-1)}| < e$ متوقف شو؛ در غیر این صورت، قرار بده $k \rightarrow k + 1$.

توجه شود که مشکل اساسی دنباله‌های تکراری (۷.۳) و (۱۲.۳) امکان وقوع صفر در مخرج آن‌ها است که باعث ناپایداری روش می‌شود. همچنین وقتی $X^{(t)}$ نزدیک به نقطه موجود و غیر بهینه A_i شود هر دو دنباله تکراری رفتارهای خوبی از خود نشان نمی‌دهند. از آنجا که روش‌هایی مبتنی بر روش وایزفیلد برای حل مساله مکان‌یابی توافقی، روش‌های پایداری نیستند بدین منظور در ادامه الگوریتم کارای گاوس- نیوتن برای حل مساله حاضر بیان می‌شود و برای بررسی درستی جواب‌های به‌دست‌آمده از الگوریتم گاوس- نیوتن، نتایج به‌دست‌آمده با نتایج الگوریتم فراابتکاری رقابت استعماری نیز مقایسه می‌شود.

۲.۴.۳ الگوریتم گاوس- نیوتن

با توجه به آن‌چه در فصل دوم بیان شد، با قرار دادن $\varphi_i(X) = \sqrt{w_i} (d'(X, A_i) - R_i)$ برای $i = 1, 2, \dots, n$ مساله هموار مکان‌یابی توافقی (۸.۳) به فرم یک مساله کمترین مربعات در می‌آید:

$$\min f_h(X) = \sum_{i=1}^n w_i (d'(X, A_i) - R_i)^2 = \frac{1}{2} \sum_{i=1}^m \varphi_i^2(X) \quad (14.3)$$

بنابراین می‌توان الگوریتم GN را برای حل مساله (۱۴.۳) به‌صورت زیر بیان کرد:

الگوریتم ۵ گاوس- نیوتن (GN) [۱۰۵]

- ۱: $X^{(0)}$ ، یک تقریب اولیه از جواب بهینه انتخاب کن و قرار بده $k = 0$.
- ۲: مقادیر $J_k^T J_k$ و $J_k^T \varphi_k$ ، $J_k = J(X_k)$ ، $\varphi_k = \varphi(X_k)$ را محاسبه کن.
- ۳: اگر $\|J_k^T \varphi_k\| < \varepsilon_1$ ، در این صورت متوقف شو؛
- ۴: معادله $J_k^T J_k d = -J_k^T \varphi_k$ را برای $d = d_k$ حل کن.
- ۵: اگر $\|d_k\| < \varepsilon_1$ ، متوقف شو؛ در غیر این صورت،
- ۶: قرار ده $X_{k+1} = X_k + \alpha d_k$ و $k + 1 \rightarrow k$ و انتقال به گام ۲.

۳.۴.۳ الگوریتم فراابتکاری رقابت استعماری [۵۸]

در این بخش ابتدا به معرفی روش رقابت استعماری پرداخته، سپس الگوریتم آن را بیان می‌کنیم و در بخش نتایج عددی نتایج حاصل از پیاده‌سازی این الگوریتم را ارائه داده و با دیگر الگوریتم‌ها مقایسه می‌کنیم.

الگوریتم‌های بهینه‌سازی، مقادیر متغیرهای طراحی را به‌گونه‌ای تغییر می‌دهند تا تابع هدف موردنظر کمینه یا بیشینه شود. روش‌های بهینه‌سازی را می‌توان به‌طور کلی به دو گروه روش‌های گرادیانی و روش‌های جستجو تقسیم نمود. روش‌های جستجو برای مسایل حجیم و مواردی که مشتق گرفتن از تابع هدف دشوار است، کارایی زیادی دارند. به‌عنوان نمونه،

بشیری و گرمه‌ای در مقاله [۱۸] و نوتاش و همکاران در مقاله [۱۰۶] از الگوریتم‌های فراابتکاری برای حل مسایل مکان‌یابی خود استفاده کرده‌اند؛ هرچند نمی‌توان به‌صراحت در مورد برتری روش‌های تکاملی بالا نسبت به یکدیگر نظر داد و این امر طبق نظریه NFLT^{۱۴} [۷۱]، از مساله‌ای به مساله دیگر متفاوت است.

همانند دیگر الگوریتم‌های تکاملی، الگوریتم رقابت استعماری نیز با تعدادی جمعیت اولیه تصادفی که هرکدام از آنها یک کشور نامیده می‌شوند، شروع می‌شود. تعدادی از بهترین عناصر جمعیت (معادل نخبه‌ها در الگوریتم ژنتیک) به‌عنوان **امپریالیست**^{۱۵} انتخاب می‌شوند. باقیمانده جمعیت نیز به‌عنوان **مستعمره**^{۱۶} در نظر گرفته می‌شوند. استعمارگران بسته به قدرتشان، این مستعمرات را با یک روند خاص به سمت خود می‌کشند. قدرت کل هر امپراتوری، به هر دو بخش تشکیل‌دهنده آن یعنی کشور امپریالیست (به‌عنوان هسته مرکزی) و مستعمرات آن، بستگی دارد. در حالت ریاضی، این وابستگی با تعریف قدرت امپراتوری به‌صورت مجموع قدرت کشور امپریالیست، به‌اضافه درصدی از میانگین قدرت مستعمرات آن، مدل شده است. با شکل‌گیری امپراتوری‌های اولیه، رقابت امپریالیستی میان آن‌ها شروع می‌شود. هر امپراتوری که نتواند در رقابت استعماری موفق عمل کند و بر قدرت خود بیفزاید (یا حداقل از کاهش نفوذش جلوگیری کند)، از صحنه رقابت استعماری حذف خواهد شد. بنابراین بقای یک امپراتوری، وابسته به قدرت آن در جذب مستعمرات امپراتوری‌های رقیب و به‌سیطره درآوردن آن‌ها خواهد بود. در نتیجه، در جریان رقابت‌های امپریالیستی به‌تدریج بر قدرت امپراتوری‌های بزرگ‌تر افزوده شده و امپراتوری‌های ضعیف‌تر حذف خواهند شد. امپراتوری‌ها برای افزایش قدرت مجبور خواهند شد تا مستعمرات خود را نیز پیشرفت دهند. با گذشت زمان، مستعمرات از لحاظ قدرت به امپراتوری‌ها نزدیک‌تر خواهند شد و یک نوع همگرایی به دست می‌آید. حد نهایی رقابت استعماری، زمانی است که یک امپراتوری واحد در دنیا وجود داشته باشد و مستعمراتش از لحاظ موقعیت به خود کشور امپریالیست خیلی نزدیک باشند.

یک الگوریتم رقابت استعماری دارای مفاهیمی و متغیرهایی همچون شکل‌دهی امپراتوری‌های اولیه، حرکت مستعمره‌ها به سمت امپریالیست، جابجایی موقعیت مستعمره و امپریالیست، قدرت کل یک امپراتوری، رقابت استعماری، سقوط امپراتوری‌های ضعیف و همگرایی است. بنابراین، الگوریتم رقابت استعماری را می‌توان به‌صورت زیر بیان نمود:

^{۱۴} No Free Lunch Theorem

^{۱۵} Imperialist

^{۱۶} Colony

الگوریتم ۶ الگوریتم رقابت استعماری (ICA)

- ۱: چند نقطه تصادفی در فضای مساله انتخاب شود و امپراتوری‌های اولیه تشکیل شوند؛
- ۲: مستعمرات، به سمت کشور امپریالیست حرکت داده شوند؛
- ۳: اگر مستعمره‌ای در یک امپراتوری وجود داشته باشد که هزینه‌ای کمتر از امپریالیست داشته باشد، جای مستعمره و امپریالیست با هم تعویض شود؛
- ۴: قدرت کل یک امپراتوری محاسبه شود (با در نظر گرفتن هزینه امپریالیست و مستعمراتشان)؛
- ۵: یک مستعمره از ضعیف‌ترین امپراتوری انتخاب‌شده به امپراتوری که بیشترین احتمال تصاحب را دارد، اختصاص داده شود؛
- ۶: امپراتوری‌های ضعیف حذف شوند؛
- ۷: اگر تنها یک امپراتوری وجود داشت، توقف و در غیر این صورت، به گام ۲ رجوع شود.

جدول ۱.۳ بیانگر مولفه‌های تنظیم‌شده الگوریتم ICA در پایان‌نامه حاضر است. این مولفه‌ها از مساله‌ای به مساله دیگر قابل تغییر است.

جدول ۱.۳: مولفه‌های تنظیم‌شده الگوریتم (ICA)

۵۰۰	بیشترین تعداد تکرار الگوریتم (Maxit)
۵۰	تعداد کشورهای اولیه ($N_{country}$)
۲	تعداد متغیرها (N_{var})
۱۰	تعداد امپراتوری‌های اولیه (N_{imp})
۲	ضریب نزدیک‌شدن یک مستعمره به استعمارگر از جهت‌های گوناگون (β)
$U(0, 1)$	میزان انحراف یک مستعمره از مسیر مستقیم حرکت به سمت استعمارگر (θ)
۰.۱	ضریب هزینه متوسط مستعمرات (ξ)
۰.۱	احتمال انقلاب

۵.۳ نتایج عددی

در این قسمت به حل مثال‌های گوناگون از مسایل مطرح‌شده پرداخته می‌شود. مدل‌های ارائه‌شده در بخش‌های قبل با سه الگوریتم WLA، الگوریتم GN و الگوریتم ICA حل می‌شوند. برای نشان دادن درستی کارایی الگوریتم‌های پیشنهادی، ابتدا مثال‌هایی در ابعاد کوچک و با

جواب بدیهی انتخاب شده و جواب‌های الگوریتم‌ها با یکدیگر مقایسه می‌شود و سپس به حل مثال‌هایی در مقیاس بزرگ پرداخته می‌شود.

تمامی الگوریتم‌ها با استفاده از نرم افزار Matlab- R2017b و با یک رایانه همراه با پردازنده Intel core i7-4510U و مقدار حافظه ۸ گیگابایت پیاده‌سازی شده‌اند.

مثال ۱.۵.۳. فرض شود $n = 6$ باشد و برای هر $i = 1, 2, \dots, 6$ ، $w_i = 1$ باشد و این نقاط روی یک دایره به مرکز (۱،۱) و شعاع ۱ قرار داشته باشند و برای هر $i = 1, 2, \dots, 6$ ، $R_i = 1$ باشد. مختصات نقاط در جدول ۲.۳ آمده است. بدیهی است مکانی که فاصله آن تا نقاط تقاضا دقیقاً برابر با یک باشد، مرکز دایره است. بنابراین نقطه (۱،۱)، باید جواب بهینه مساله باشد. درستی این امر را می‌توان در نتایج به‌دست‌آمده از الگوریتم‌های WLA، GN و ICA مشاهده کرد.

جدول ۲.۳: مختصات نقاط تقاضا $A_i = (a_i, b_i)$ برای حالت $p = 2$ و $w_i = R_i = 1$

	A_1	A_2	A_3	A_4	A_5	A_6
a_i	۲	۱	۰	۱	$\frac{2+\sqrt{2}}{2}$	$\frac{2+\sqrt{2}}{2}$
b_i	۱	۲	۱	۰	$\frac{2+\sqrt{2}}{2}$	$\frac{2-\sqrt{2}}{2}$

در مثال بالا، هر سه الگوریتم WLA، GN و ICA به جواب $X = (1, 1)$ رسیدند. در مثال بالا برای الگوریتم WLA و GN، $\varepsilon_1 = 10^{-4}$ در نظر گرفته شده است.

مثال ۲.۵.۳. جداول ۳.۳ و ۴.۳ حاوی نقاط اولیه و شعاع‌های توافقی و وزن نقاط تقاضا است. در هر یک از جداول ۳.۳ و ۴.۳ سه نوع شعاع توافقی متفاوت برای هر نقطه تقاضا در نظر گرفته شده است. در جداول ۵.۳ و ۶.۳ به ترتیب، نتایج به‌دست‌آمده از حل داده‌های جداول ۳.۳ و ۴.۳ برای تابع هدف مساله (۸.۳) آمده است و نتایج آن با روش BSS که در مرجع [۵۳] آمده است، مقایسه شده است شکل ۲.۳، دو نمای گوناگون از تابع هدف مساله (۸.۳) با داده‌های جدول ۳.۳ را برای شعاع توافقی I نشان می‌دهد. نمودارهای ۴.۳ و ۳.۳ نیز نحوه کاهش گرادیان تابع هدف مساله را برای داده‌های جدول ۳.۳ و در حالت شعاع توافقی II، برای هریک از الگوریتم‌های ICA و GN نمایش داده است. جدول ۶.۳ نیز نتایج به‌دست‌آمده از داده‌های جدول ۴.۳ نسبت به نرم‌های گوناگون را بیان می‌کند. جدول ۷.۳ نتایج به‌دست‌آمده از حل مساله با الگوریتم‌های ارائه‌شده در این فصل را نسبت به مقادیر گوناگون n و p نشان می‌دهد. توجه شود از آن‌جا که روش‌های گاوس- نیوتن و رقابت استعماری نسبت به روش شبه وایزفیلد عملکرد بهتری از خود نشان داده‌اند، فقط نتایج این دو روش در جداول مربوطه گنجانده شده است.

جدول ۳.۳: داده‌های یک مساله با ۴ نقطه برای حالت $p = 2$

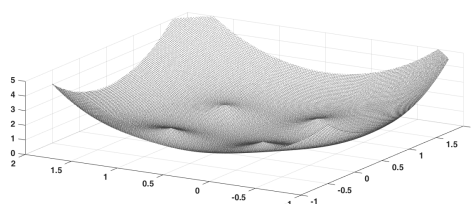
(a_i, b_i, w_i)	شعاع توافقی I	شعاع توافقی II	شعاع توافقی III
$(0, 0, 1)$	۱	۱	۲
$(1, 0, 1)$	۱	۲	۲
$(0, 1, 1)$	۱	۱	۲
$(1, 1, 1)$	۱	۲	۲

جدول ۴.۳: داده‌های یک مساله با ۳۰ نقطه برای $R_i = 1$

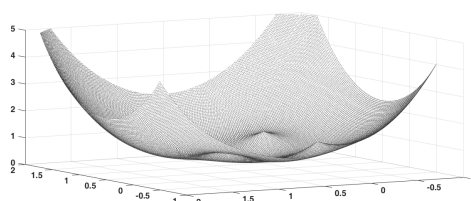
(a_i, b_i, w_i)	(a_i, b_i, w_i)	(a_i, b_i, w_i)	(a_i, b_i, w_i)	(a_i, b_i, w_i)	(a_i, b_i, w_i)
$(1, 3, 3)$	$(7, 15, 2)$	$(14, 15, 1)$	$(3, 2, 1)$	$(10, 8, 1)$	$(3, 10, 2)$
$(1, 4, 2)$	$(8, 3, 1)$	$(14, 3, 1)$	$(4, 6, 1)$	$(10, 10, 3)$	$(9, 11, 2)$
$(2, 15, 1)$	$(8, 6, 3)$	$(14, 1, 2)$	$(4, 3, 2)$	$(11, 4, 2)$	$(15, 15, 2)$
$(2, 4, 3)$	$(8, 5, 1)$	$(15, 8, 3)$	$(6, 8, 3)$	$(11, 13, 3)$	$(7, 14, 2)$
$(3, 6, 2)$	$(8, 2, 1)$	$(15, 10, 3)$	$(6, 11, 1)$	$(13, 3, 1)$	$(13, 7, 3)$

جدول ۵.۳: نتایج به‌دست‌آمده از داده‌های جدول ۳.۳ برای $p = 2$

		شعاع توافقی I	شعاع توافقی II	شعاع توافقی III
BSSS	X	$(0/49, 0/51)$	$(-0/90, 0/51)$	$(-1/42, 0/51)$
	f	۰.۳۴	۰.۰۰	۰.۹۳
GN	X	$(0/50, 0/50)$	$(-0/9049, 0/50)$	$(-1/4228, 0/50)$
	f	۰.۳۴۳۱	۰.۰۰۴۲	۰.۹۳۳۰
ICA	X	$(0/499, 0/50)$	$(-0/9048, 0/50)$	$(0/50, -1/4151)$
	f	۰.۳۴۳۱۱	۰.۰۰۴۱	۰.۹۲۷۶

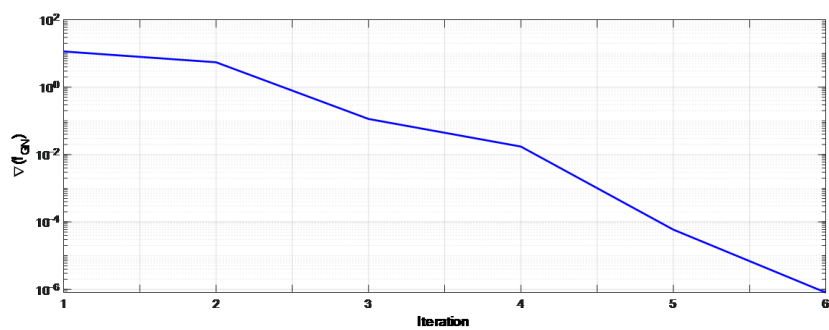


(آ)

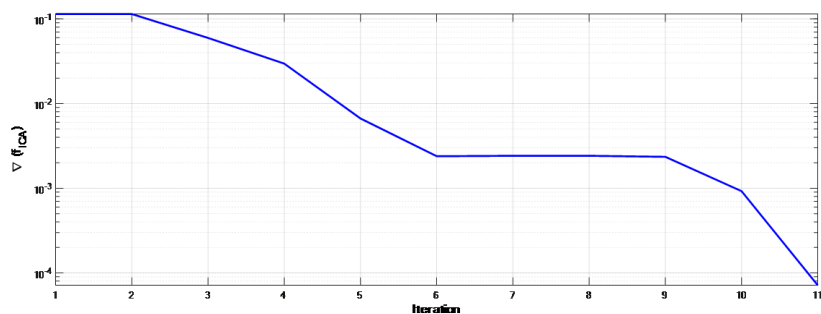


(ب)

شکل ۲.۳: دو نمای گوناگون از تابع هدف مساله (۸.۲) با داده‌های جدول ۳.۳ و شعاع توافقی I



شکل ۳.۳: نمودار کاهشی گرادیان تابع هدف الگوریتم GN برای مثال جدول ۳.۳ با شعاع توافقی II



شکل ۴.۳: نمودار کاهشی گرادیان تابع هدف الگوریتم ICA برای مثال جدول ۵.۳ با شعاع توافقی II

جدول ۶.۳: نتایج به دست آمده از داده‌های جدول ۴.۳ نسبت به مقادیر مختلف p

p	f_{GN}	f_{ICA}	زمان اجرای GN	زمان اجرای ICA
۱	۳۱۵۶.۰	۳۱۵۶.۰۲	۷.۲۷	۹۶.۵۵
۱.۵	۲۰۳۳.۷۰	۲۰۳۳.۷۱	۱۸.۵۱	۶۴.۸۲
۲	۱۶۶۸.۱	۱۶۶۸.۱۲	۹.۳۳	۹۵.۶۸
۳	۱۴۰۴.۱	۱۴۰۴.۱۴	۹.۸۱	۵۱.۸۴
۴	۱۳۰۵.۷	۱۳۰۵.۶۹	۱۳.۲۷	۶۱.۰۴
۵	۱۲۵۶.۸	۱۲۵۶.۷	۱۰.۹۶	۷۴.۳۷
۱۰	۱۱۸۵.۰	۱۱۸۴.۹	۸.۲۲	۵۱.۶۳

جدول ۷.۳: نتایج حاصله از الگوریتم‌های GN و ICA نسبت به مقدار مختلف n و p

n	p	f_{GN}	زمان اجرای GN	f_{ICA}	زمان اجرای ICA
۱۰۰	۱.۵	۸۴۵۷۰	۱۹.۷۰	۸۴۵۶۹.۹	۱۶۶.۴۰
۱۰۰	۲	۶۸۹۹۵	۱۳.۶۶	۶۸۹۹۵.۲	۱۰۴.۱۲
۱۰۰	۳	۵۸۰۵۷	۱۹.۳۹	۵۸۰۵۶.۷	۱۲۵.۶۹
۲۰۰	۱.۵	۲۰۴۰۵۰	۳۷.۳۱	۲۰۴۰۴۶	۳۶۵.۹۷
۲۰۰	۲	۱۶۶۲۵۰	۲۱.۸۳	۱۶۶۲۵۴	۱۵۱.۳۱
۲۰۰	۳	۱۳۹۹۴۰	۳۱.۶۷	۱۳۹۹۴۰.۱	۲۵۰.۴۰
۳۰۰	۱.۵	۲۸۸۷۰۰	۳۸.۲۷	۲۸۸۶۹۵.۴	۴۵۷.۱۲
۳۰۰	۲	۲۳۴۳۱۰	۲۷.۴۷	۲۳۴۳۱۴.۳	۳۵۳.۸۳
۳۰۰	۳	۱۹۵۸۶۰	۳۷.۲۲	۱۹۵۸۵۵.۷	۴۲۹.۱۶
۵۰۰	۱.۵	۴۶۶۹۲۰	۶۴.۰۹	۴۶۶۹۱۶.۴	۵۴۲.۱۰
۵۰۰	۲	۳۸۱۷۸۰	۴۷.۵۴	۳۸۱۷۷۹.۲	۵۵۲.۱۰
۵۰۰	۳	۳۲۱۸۳۰	۶۳.۶۲	۳۲۱۱۷۹.۱	۵۷۳.۲۳
۱۰۰۰	۱.۵	۹۸۲۶۳۰	۱۶۱.۳۵	۹۸۲۶۳۱.۳	۷۸۵.۱۴
۱۰۰۰	۲	۷۹۷۱۵۰	۱۳۲.۰۴	۷۹۱۴۸.۴	۶۵۴.۷۴
۱۰۰۰	۳	۶۶۶۱۷۰	۱۶۹.۹۳	۶۶۶۱۹۸.۴	۷۲۱.۱۵

نتایج به‌دست‌آمده از حل مثال‌هایی با ابعاد بزرگ و کوچک نشان داده است که جواب‌های به‌دست‌آمده از الگوریتم‌های GN و ICA تطابق بسیار خوبی با یکدیگر دارند؛ به‌طوری‌که گرادیان تابع هدف برای یک مثال خاص برای الگوریتم ICA در ۱۱ تکرار (نمودار ۴.۳) و برای الگوریتم GN در ۶ تکرار (نمودار ۳.۳) به سمت صفر میل کرده است. بنابراین با توجه به اینکه زمان اجرای الگوریتم GN نسبت به الگوریتم فراابتکاری ICA به‌مراتب کمتر است، لذا این الگوریتم گرادینانی از لحاظ زمانی نسبت به الگوریتم ICA از برتری بیشتری برخوردار است.

فصل ۴

مسایل مکان‌یابی و بر توافقی تحت تابع زیان نامتقارن

در بسیاری از موقعیت‌های زندگی واقعی، خطای مثبت و خطای منفی با اندازه‌های یکسان، اغلب مفاهیم متفاوت اقتصادی و مادی دارند، به عبارتی خطای مثبت ممکن است جدی‌تر یا مهم‌تر از خطای منفی باشد و یا بالعکس، بنابراین در این صورت استفاده از توابع زیان متقارن مانند تابع زیان مربع خطا یا تابع زیان قدر مطلق خطا، که قادر به تفکیک این موضوع نمی‌باشند، نامناسب هستند.

به عنوان یک مثال عینی می‌توان به این مورد اشاره نمود که در ساخت و ساز سد، کم برآورد حداکثر سطح آب، بسیار مهم‌تر و جدی‌تر از بیش برآورد آن است زیرا ممکن است در اثر کم برآورد سطح آب در یک طوفان سیل آسا آب پشت سد از دیواره آن بالا آمده و از آن خارج شود و خسارت‌های جانی و مالی زیادی را به همراه بیاورد در حالی که بیش برآورد حداکثر سطح آب تنها ممکن است خسارت ناشی از هزینه‌های ساخت و ساز را در پی داشته باشد. ^۱ واریان [۱۳۴] در یک مطالعه کاربردی در ارزیابی مالکیت املاک و مستقالات، تابع زیان بسیار مفید نامتقارن ^۲ لاینکس را معرفی نمود. این تابع در یک طرف صفر، تقریباً به صورت نمایی افزایش می‌یابد درحالی که در سمت دیگر آن به صورت خطی رفتار می‌کند.

^۱ Varian

^۲ linex

تعریف ۱.۰.۴ (تابع زیان نامتقارن لاینکس [۱۳۴]). فرض کنید که δ برآوردگر پارامتر θ باشد و $\Delta = \theta - \delta$. در این صورت تابع زیان محدب لاینکس توسط واریان به صورت زیر معرفی شد.

$$E(\Delta) = be^{a\Delta} - c\Delta - b, \quad a, c \neq 0, b > 0 \quad (1.4)$$

در رابطه بالا مشاهده می‌شود که $E(0) = 0$ است. همچنین برای وجود مینیمم در $\Delta = 0$ کافی است $ab = c$ باشد. بنابراین با این تبدیل می‌توانیم رابطه (۱.۴) را به صورت زیر بازنویسی نمائیم.

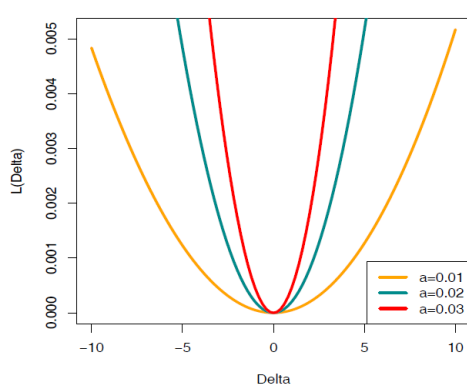
$$E(\Delta) = b[e^{a\Delta} - a\Delta - 1], \quad a \neq 0, b > 0 \quad (2.4)$$

شکل ۱.۴، نمودار تابع زیان $E(\Delta)$ را به‌ازای مقادیر متفاوتی از a و با فرض $b = 1$ نشان می‌دهد. همان‌طور که در شکل ۱.۴ مشاهده می‌شود به‌ازای $a = 1$ تابع کاملاً نامتقارن و ارزش بیش برآورد نسبت به کم برآورد بیشتر است و برای $a = -1$ عکس آن برقرار است. همچنین اگر بسط نمایی تابع $e^{a\Delta}$ را در نظر بگیریم در این صورت داریم، $e^{a\Delta} \simeq 1 + a\Delta + \frac{a^2\Delta^2}{2}$ ، بنابراین برای مقادیر کوچک $|a|$ ، تابع زیان لاینکس تقریباً متقارن است و اختلاف چندانی با تابع زیان مربع خطا ندارد. این امر را می‌توان در شکل ۱.۴ مشاهده نمود.

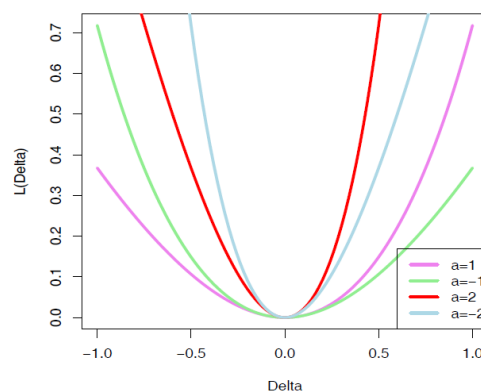
آرشی [۱۰]، تعمیمی از تابع زیان لاینکس را به صورت زیر بیان نمود:

$$E(\Delta) = \sum_{i=1}^k b_i [e^{a_i\Delta_i} - a_i\Delta_i - 1], a_i \neq 0, b_i > 0, \forall i = 1, \dots, k \quad (3.4)$$

که در آن $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_k)^T$ است. این تابع زیان محدب است و $E'(0) = 0$ ، $E(0) = 0$ و $\Delta = 0$ مینیمم این تابع است.



(ب)



(ا)

شکل ۱.۴: نمودار تابع زیان لاینکس نسبت به مقادیر مختلف a در حالت $b = 1$.

۱.۴ مساله مکان‌یابی وبر توافقی تحت تابع زیان نامتقارن لاینکس

از آنجایی که توابع زیان مربعات خطا و قدرمطلق توابعی نامحدب هستند و از طرف دیگر، این توابع زیان متقارن بوده و قادر به ارزش‌گذاری بین خطاهای مثبت و منفی با اندازه‌هایی یکسان نیستند، لذا در این قسمت تابع زیان لاینکس که هم تابعی محدب است و هم نامتقارن را به عنوان تابع زیان برای مساله وبر توافقی در نظر می‌گیریم. بنابراین یک مساله وبر توافقی تحت تابع زیان نامتقارن لاینکس و با نرم ℓ_p ، به صورت زیر مدل‌بندی می‌شود:

$$f_L(X) = \sum_{i=1}^n w_i \cdot \left(e^{a \cdot (d(X, A_i) - R_i)} - a \cdot (d(X, A_i) - R_i) - 1 \right) \quad (۴.۴)$$

در رابطه (۴.۴) با در نظر گرفتن $a = \pm 1$ ، به یک تابع کاملاً نامتقارن می‌رسیم درحالی‌که اگر مقدار $|a|$ را کوچک در نظر بگیریم تابع زیان لاینکس رفتاری مشابه تابع زیان متقارن مربعات خطا دارد. بنابراین با در نظر گرفتن مقادیر مختلف می‌توان رفتارهای مختلف توابع متقارن و نامتقارن را شبیه‌سازی نمود.

قضیه ۱.۱.۴. جواب بهینه مساله مکان‌یابی وبر توافقی تحت تابع زیان نامتقارن لاینکس و با نرم ℓ_p ، در پوسته گسترش یافته مستطیلی نقاط $A_i = (a_i, b_i)$ قرار دارد.

برهان. فرض کنید $\bar{X} = (\bar{x}, \bar{y})$ ، خارج از پوسته مستطیلی نقاط RH_1, RH_2, RH_3 و RH_4 قرار داشته باشد. در این صورت، چهار حالت ممکن است رخ دهد: $\bar{x} < a_{min}$ ، $\bar{x} > a_{max}$ ، $\bar{y} > b_{max}$ و $\bar{y} < b_{min}$. به سادگی می‌توان نشان داد که در هر حالت $\bar{X}$ بهینه نیست. در اینجا اثبات حالت اول را ارائه می‌دهیم. اثبات حالات دیگر مشابه است.

فرض کنید $\bar{x} > a_{max}$ ، در این صورت، قرار دهید $X' = (a_{max}, \bar{y})$. برای هر $i = 1, \dots, n$ داریم

$$d(\bar{X}, A_i) = (|\bar{x} - a_i|^p + |\bar{y} - b_i|^p)^{\frac{1}{p}} > (|a_{max} - a_i|^p + |\bar{y} - b_i|^p)^{\frac{1}{p}} = d(X', A_i) > R_i$$

بنابراین،

$$\left((|\bar{x} - a_i|^p + |\bar{y} - b_i|^p)^{\frac{1}{p}} - R_i \right) > \left((|a_{max} - a_i|^p + |\bar{y} - b_i|^p)^{\frac{1}{p}} - R_i \right) > 0$$

با توجه به مثبت بودن ضرائب w_i و صعودی بودن تابع لاینکس در قسمت مثبت نمودار نتیجه می‌شود:

$$w_i \cdot \left(e^{a \cdot (d(\bar{X}, A_i) - R_i)} - a \cdot (d(\bar{X}, A_i) - R_i) - 1 \right) > w_i \cdot \left(e^{a \cdot (d(X', A_i) - R_i)} - a \cdot (d(X', A_i) - R_i) - 1 \right)$$

لذا می‌توان نتیجه گرفت $f_L(\bar{X}) > f_L(X')$ بنابراین $\bar{X}$ نمی‌تواند جواب بهینه باشد. $\square$

حال در پی جواب بهینه مساله (۴.۴) هستیم. چون تابع لاینکس محدب است و برای هر $w_i \geq 0, i = 1, \dots, n$ بنابراین لم زیر حاصل می‌شود.

لم ۱.۱.۴. تابع هدف مساله (۴.۴) محدب است.

توجه داریم که همان‌طور که فتحعلی و همکاران [۵۳] نشان داده‌اند، مساله (۴.۳) با نرم اقلیدسی محدب نیست و ممکن است جواب بهینه موضعی، بهینه سراسری نباشد. اما، برای تابع لاینکس جواب بهینه موضعی، بهینه سراسری است.

۱.۱.۴ روش شبه- وایزفیلد

مطابق آنچه در فصل قبل بیان کردیم، به منظور مشتق‌پذیری در نقاط موجود، از تقریب هموار f_L^h استفاده می‌کنیم.

$$f_L^h(X) = \sum_{i=1}^n w_i \cdot \left(e^{a \cdot (d'(X, A_i) - R_i)} - a \cdot (d'(\bar{X}, A_i) - R_i) - 1 \right) \quad (۵.۴)$$

که در آن

$$d'(X, A_i) = \left(\left((x - a_i)^2 + \epsilon \right)^{\frac{p}{2}} + \left((y - b_i)^2 + \epsilon \right)^{\frac{p}{2}} \right)^{\frac{1}{p}}$$

و ϵ یک عدد مثبت کوچک است.

لم ۲.۱.۴.

$$\lim_{\epsilon \rightarrow 0} |f_L^h(X) - f_L(X)| = 0$$

برهان. با استفاده از نامساوی مینکوفسکی داریم:

$$\begin{aligned} & \left| w_i \cdot \left(e^{a \cdot (d'(\bar{X}, A_i) - R_i)} - a \cdot (d'(\bar{X}, A_i) - R_i) - 1 \right) - w_i \cdot \left(e^{a \cdot (d(\bar{X}, A_i) - R_i)} - a \cdot (d(\bar{X}, A_i) - R_i) - 1 \right) \right| \\ &= \left| w_i \cdot \left(e^{a \cdot (d'(\bar{X}, A_i) - R_i)} - e^{a \cdot (d(\bar{X}, A_i) - R_i)} \right) - a \cdot w_i (d'(\bar{X}, A_i) - d(\bar{X}, A_i)) \right| \\ &\leq w_i \left(\left| e^{a \cdot (d'(\bar{X}, A_i) - R_i)} - e^{a \cdot (d(\bar{X}, A_i) - R_i)} \right| + |a| |d'(\bar{X}, A_i) - d(\bar{X}, A_i)| \right) \\ &\leq w_i \left(\left| e^{a \cdot (d(\bar{X}, A_i) - R_i)} \left(e^{(a \cdot (d'(\bar{X}, A_i) - d(\bar{X}, A_i)))} - 1 \right) \right| + |a| |d'(\bar{X}, A_i) - d(\bar{X}, A_i)| \right) \end{aligned}$$

بنابراین،

$$\left| f_L^h(\bar{X}) - f_L(\bar{X}) \right| \leq \sum_{i=1}^n w_i \left(\left| e^{a \cdot (d(\bar{X}, A_i) - R_i)} \left(e^{(a \cdot (d'(\bar{X}, A_i) - d(\bar{X}, A_i)))} - 1 \right) \right| + |a| |d'(\bar{X}, A_i) - d(\bar{X}, A_i)| \right)$$

در نتیجه،

$$\lim_{\epsilon \rightarrow 0} |f_L^h(X) - f_L(X)| = 0$$

□

حال، شرط لازم بهینگی برای $f_L^h(X)$ را به کار می‌گیریم.

(۶.۴)

$$\frac{\partial f_L^h}{\partial X_K} = a \sum_{i=1}^n w_i \left(\frac{d'''(X_k, a_{ik}) \cdot (e^{a \cdot (d'(X, A_i) - R_i)} - 1)}{d''(X, A_i)} (X_k - a_{ik}) \right) = 0, \quad k = 1, 2$$

که

$$\begin{cases} d''(X, A_i) = \left(\left((x - a_i)^2 + \epsilon \right)^{\frac{p}{2}} + \left((y - b_i)^2 + \epsilon \right)^{\frac{p}{2}} \right)^{\frac{p-1}{p}} \\ d'''(X_k, a_{ik}) = \left((x - a_{ik})^2 + \epsilon \right)^{\frac{p-2}{2}}, \quad k = 1, 2 \end{cases}$$

بنابراین، با شروع از نقطه اولیه $X^\circ = (x^\circ, y^\circ)$ ، قاعده تکراری زیر را به کار می‌گیریم:

$$X_k^{(t+1)} = \frac{\sum_{i=1}^n \left(\frac{w_i \cdot d'''(X_k, a_{ik}^{(t)}) \cdot (e^{a \cdot (d'(X^t, A_i) - R_i)} - 1)}{d''(X, A_i^{(t)})} a_{ik} \right)}{\sum_{i=1}^n \left(\frac{w_i \cdot d'''(X_k, a_{ik}^{(t)}) \cdot (e^{a \cdot (d'(X^t, A_i) - R_i)} - 1)}{d''(X, A_i^{(t)})} \right)}, \quad k = 1, 2 \quad (7.4)$$

ملاحظه ۱.۱.۴. به طور مشابه، ملاحظه ۱.۴.۳ و لم‌های ۲.۴.۳، ۳.۴.۳ و ۴.۴.۳ فصل قبل برای روش تکراری فوق نیز برقرار است.

در واقع، روش شبه- وایزفیلد روش تکراری نقطه ثابت است که در حالت کلی دارای سرعت همگرایی خطی است. با توجه به مشکلات مطرح شده برای این روش در فصل قبل به منظور حل مساله (۵.۴)، در ادامه روش‌های مبتنی بر BFGS را به کار می‌گیریم که از روش شبه- وایزفیلد سریع‌تر هستند.

۲.۱.۴ روش BFGS و نسخه‌های اصلاح‌شده آن

بنا به آنچه در فصل دوم بیان شد، روش BFGS یک روش متداول شبه نیوتن است که در هر تکرار به جای استفاده از هسی، تقریبی از هسی را به کار می‌گیرد. این روش خصوصیات همگرایی مطلوبی نسبت به روش‌های نیوتن دارد [۱۰۵].

در ادبیات موضوع، روش‌های اصلاح شده زیادی برای روش BFGS وجود دارند که به حل مسایل بهینه‌سازی نامقید می‌پردازند و مورد توجه بسیاری از محققان قرار گرفته است، بدین منظور از چند نسخه اصلاح‌شده جدید برای حل مساله (۵.۴) استفاده می‌کنیم. در بخش بعدی، شش نسخه اصلاح‌شده روش BFGS را به کار می‌گیریم که این روش‌ها در جدول ۱.۴ آورده شده‌اند.

جدول ۱.۴: نسخه‌های اصلاح شده روش BFGS و نویسندگان آنها

نویسندگان	روش‌ها
اندري ^۳ [۹]	SBFGS: an adaptive Scaled BFGS method
اندري [۸]	TSBFGS: A double parameter Scaled BFGS method
یوان و زنگین ^۴ [۱۴۳]	MBFGS: a Modified BFGS method
بابایی کفاکی [۱۱]	MMLSBFGS: A modified scaling parameter for the memoryless BFGS updating formula
بابایی کفاکی [۱۲]	TMLSBFGS: Two-parameter Scaled Memoryless BFGS methods
ارن و لوئنبگر ^۵ [۱۰۷]	SMMLSBFGS: Self-scaling variable metric (ssvm) algorithms

۲.۴ نتایج عددی

در این بخش، به مقایسه نتایج عددی حاصل از الگوریتم‌های BFGS و WLA با ICA و GN برای مسایل مختلف می‌پردازیم. تمامی الگوریتم‌ها با استفاده از نرم افزار Matlab- R2017b و با یک رایانه همراه با پردازنده Intel core i7-4510U و مقدار حافظه ۸ گیگابایت پیاده‌سازی شده‌اند. توجه شود که پارامترهای ρ و σ در جستجوی خطی ولف به ترتیب برابر 10^{-4} و 10^{-1} در نظر گرفته شده‌اند و تقریب آغازین ماتریس هسی نیز برابر ماتریس همانی فرض شده است.

ابتدا مساله‌ای با $n = 30$ نقطه در نظر بگیرید که اطلاعات تکمیلی آن در جدول ۲.۴ داده شده است. در حالت $p = 2$ همه روش‌ها به نقاط $(8.34, 8.08)$ ، $(8.30, 8.74)$ و $(8.34, 8.09)$ با مقادیر بهینه ۴۵۱۲۱، ۱۲۸۱۶ و ۵۹۲۸ به ترتیب برای سه شعاع توافقی همگرا می‌شوند. جدول ۳.۴ زمان اجرای الگوریتم‌های مذکور را نشان می‌دهد.

همان‌طور که نتایج نشان می‌دهند، روش‌های BFGS از دیگر روش‌ها سریع‌تر هستند. جدول ۴.۴ نتایج حاصله از روش BFGS برای نرم‌های مختلف و شعاع توافقی دوم را نشان می‌دهد.

^۳ Andrei

^۴ Yuan and Zengxin

جدول ۲.۴: مختصات، وزن‌ها و شعاع توافقی نقاط موجود

شعاع توافقی (x, y, w)	شعاع توافقی (x, y, w)	شعاع توافقی (x, y, w)
(۱, ۳, ۳)	(۱, ۲, ۳)	(۱۴, ۱۵, ۱)
(۱, ۴, ۲)	(۱, ۲, ۳)	(۱۴, ۳, ۱)
(۲, ۱۵, ۱)	(۱, ۳, ۳)	(۱۴, ۱, ۲)
(۱, ۳, ۳)	(۱, ۲, ۳)	(۱۴, ۱۵, ۱)
(۲, ۴, ۳)	(۱, ۲, ۳)	(۱۵, ۸, ۳)
(۳, ۶, ۲)	(۱, ۲, ۳)	(۱۵, ۱۰, ۳)
(۷, ۱۵, ۲)	(۱, ۲, ۳)	(۳, ۲, ۱)
(۸, ۳, ۱)	(۱, ۲, ۳)	(۴, ۶, ۱)
(۸, ۶, ۳)	(۱, ۲, ۳)	(۴, ۳, ۲)
(۸, ۵, ۱)	(۱, ۲, ۳)	(۶, ۸, ۳)
(۸, ۲, ۱)	(۱, ۳, ۳)	(۶, ۱۱, ۱)
(۱۰, ۸, ۱)	(۱, ۱, ۳)	(۱۰, ۸, ۱)
(۱۰, ۱۰, ۳)	(۱, ۱, ۳)	(۱۰, ۱۰, ۳)
(۱۱, ۴, ۲)	(۱, ۲, ۳)	(۱۱, ۴, ۲)
(۱۰, ۸, ۱)	(۱, ۱, ۳)	(۱۰, ۸, ۱)
(۱۱, ۱۳, ۳)	(۱, ۱, ۳)	(۱۱, ۱۳, ۳)
(۱۳, ۳, ۱)	(۱, ۲, ۳)	(۱۳, ۳, ۱)
(۳, ۱۰, ۲)	(۱, ۲, ۳)	(۳, ۱۰, ۲)
(۹, ۱۱, ۲)	(۱, ۲, ۳)	(۹, ۱۱, ۲)
(۱۵, ۱۵, ۲)	(۱, ۲, ۳)	(۱۵, ۱۵, ۲)
(۷, ۱۴, ۲)	(۱, ۲, ۳)	(۷, ۱۴, ۲)
(۱۳, ۷, ۳)	(۱, ۳, ۳)	(۱۳, ۷, ۳)

جدول ۳.۴: زمان اجرای روش‌ها

نام روش	زمان اجرا بر حسب ثانیه
ICA	۱۴/۲۵
GN	۴/۳۰
WLA	۱۶۰/۲۱
BFGS	۲/۳۱
SBFGS	۲/۲۹
TSBFGS	۲/۵۱
MBFGS	۲/۰۵
MMLSBFGS	۱/۹۵
TMLSBFGS	۱/۹۶
SMMLSBFGS	۲/۰۱

جدول ۴.۴: نتایج حاصله از BFGS برای مقادیر مختلف p و شعاع توافقی دوم

p	X	$f(X)$
۱	(۸/۵۸, ۷/۷۲)	۳۸۶۵۴۰
۱/۵	(۸/۴۲, ۷/۶۸)	۳۲۹۵۰
۲	(۸/۳۰, ۷/۷۴)	۱۲۸۱۶
۵	(۸/۱۸, ۸/۱۲)	۴۱۶۲
۱۰	(۸/۱۳, ۸/۲۶)	۳۴۸۴

به علاوه، الگوریتم‌های مذکور را روی مثال‌هایی با تعداد نقاط ۱۰۰، ۲۰۰ و ۵۰۰ آزمون کرده‌ایم که مختصات، وزن‌ها و شعاع توافقی آن‌ها به صورت تصادفی تولید شده‌اند. نقاط

بهینه به دست آمده توسط این الگوریتم‌ها در جدول ۵.۴ ارائه شده است. به دلیل این که همه روش‌ها به جواب‌های نزدیک به هم دست یافته‌اند، زمان اجرا و تعداد تکرارها آن‌ها را در جدول ۶.۴ مقایسه کرده‌ایم. چون زمان اجرای روش‌های ICA و WLA بسیار زیاد است و روش‌های BFGS خیلی سریع‌تر از روش GN است، فقط نتایج روش‌های BFGS را مقایسه کرده‌ایم.

جدول ۵.۴: نتایج حاصله از مثال‌های مقیاس بزرگ برای مقادیر مختلف p

n	p	$X = (x, y)$
۱۰۰	۱/۵	(۳۱/۳۸, ۳۳/۸۹)
۱۰۰	۲	(۳۱/۵۳, ۳۳/۶۵)
۱۰۰	۳	(۳۱/۹۶, ۳۲/۷۸)
۲۰۰	۱/۵	(۳۰/۹۴, ۳۱/۹۴)
۲۰۰	۲	(۳۰/۹۴, ۳۲/۰۲)
۲۰۰	۳	(۳۱/۰۲, ۳۱/۹۶)
۵۰۰	۱/۵	(۳۲/۲۳, ۳۱/۲۲)
۵۰۰	۲	(۳۲/۲۲, ۳۱/۱۶)
۵۰۰	۳	(۳۲/۱۱, ۳۱/۰۶)

جدول ۶.۴: زمان اجرای چند تکرار برای مثال‌های در مقیاس بزرگ برای مقادیر مختلف p

روش	n	$p = 1/5$		$p = 2$		$p = 3$	
		تعداد تکرارها زمان اجرا		تعداد تکرارها زمان اجرا		تعداد تکرارها زمان اجرا	
BFGS	۱۰۰	۱۱/۱۸	۸	۶/۲۰	۷	۸/۳۱	۶
	۲۰۰	۱۸/۷۳	۵	۱۱/۰۴	۶	۱۷/۱۶	۵
	۵۰۰	۴۱/۰۲	۶	۲۷/۲۸	۵	۴۴/۷۳	۶
SBFGS	۱۰۰	۱۱/۱۱	۸	۶/۱۸	۷	۸/۴۱	۶
	۲۰۰	۱۸/۱۰	۵	۱۲/۵۶	۷	۱۷/۱۵	۵
	۵۰۰	۴۰/۹۱	۶	۲۷/۱۳	۵	۴۴/۶۳	۶
TSBFGS	۱۰۰	۱۱/۵۹	۹	۷/۷۴	۹	۹/۳۲	۶
	۲۰۰	۲۱/۹۵	۶	۱۳/۷۴	۷	۲۱/۸۴	۸
	۵۰۰	۴۵/۱۹	۸	۳۰/۲۶	۶	۴۱/۲۵	۵
MBFGS	۱۰۰	۹/۸۷	۷	۵/۳۳	۴	۹/۳۹	۵
	۲۰۰	۱۷/۹۲	۶	۱۳/۵۴	۷	۱۷/۲۸	۵
	۵۰۰	۳۴/۷۳	۴	۲۷/۲۰	۶	۴۲/۳۹	۵
MMLSBFGS	۱۰۰	۷/۹۴	۶	۶/۰۳	۷	۸/۲۱	۶
	۲۰۰	۱۶/۷۴	۵	۱۲/۰۳	۵	۱۳/۵۴	۴
	۵۰۰	۳۵/۵۰	۵	۲۷/۲۳	۵	۳۶/۵۲	۴
TMLSBFGS	۱۰۰	۹/۳۹	۷	۶/۴۳	۷	۹/۱۰	۶
	۲۰۰	۱۶/۶۱	۵	۱۲/۰۶	۵	۱۶/۴۶	۵
	۵۰۰	۳۵/۶۵	۵	۲۷/۴۳	۵	۳۹/۰۶	۵
SMMLSBFGS	۱۰۰	۹/۳۸	۷	۶/۴۵	۷	۹/۱۰	۶
	۲۰۰	۱۶/۹۵	۵	۱۲/۳۲	۵	۱۶/۵۱	۵
	۵۰۰	۳۵/۸۲	۵	۲۷/۲۷	۵	۳۸/۸۰	۵

مقایسه نتایج عددی به دست آمده با استفاده از معیار دولان – موره

یکی از معیارهایی که اخیراً به طور گسترده برای مقایسه کارایی عددی الگوریتم‌های بهینه‌سازی مورد توجه قرار گرفته است، معیار ارائه شده توسط دولان و موره^۵ [۴۱] است. در ادامه، به صورت مختصر معیار دولان و موره را بررسی می‌کنیم.

فرض کنید که می‌خواهیم کارایی یک مجموعه S از روش‌های حل را روی یک مجموعه P از مسایل آزمون مقایسه کنیم. فرض می‌کنیم که n_P و n_S به ترتیب نشان دهنده تعداد روش‌های حل و تعداد مسایل آزمون باشد. حال، برای هر مساله $p \in P$ و روش $s \in S$ تعریف می‌کنیم: $t_{p,s}$: زمان محاسبه یا تعداد محاسبات مورد نیاز برای حل مساله p با روش s . برای مقایسه عملکرد روش s روی مساله p با عملکرد بهترین روش برای مساله p ، تعریف زیر را ارائه می‌کنیم:

$$r_{p,s} = \frac{t_{p,s}}{\min \{t_{p,s} : s \in S\}},$$

و قرار می‌دهیم $r_{p,s} = r_M$ اگر و تنها اگر روش s نتواند مساله p را حل کند. r_M به صورت زیر انتخاب می‌شود:

$$\forall p, s \quad r_M > r_{p,s}.$$

دولان و موره نشان دادند که انتخاب r_M در مقایسه کارایی تأثیرگذار نیست. آن‌ها همچنین به منظور نشان دادن کارایی الگوریتم s روی مجموعه مسایل آزمونی P ، تعریف زیر را ارائه دادند:

$$\rho_s(\tau) = \frac{1}{n_P} \{p \in P : r_{p,s} \leq \tau\},$$

که به ازای مقادیر مختلف $\tau \geq 1$ ، همان احتمال کمتر بودن $r_{p,s}$ از τ است. تابع ρ_s یک تابع توزیع تجمعی برای $r_{p,s}$ است. این تابع، تابعی افزایشی و قطعه به قطعه ثابت است که در نقاط انفصال تابع از راست پیوسته است.

نمودار تابع ρ_s اکثر خواص اجرای روش‌ها را نشان می‌دهد، به خصوص اگر تعداد اعضای مجموعه P به اندازه کافی بزرگ باشد. از این نمودار برای مقایسه عملکرد روش‌های مختلف می‌توان بهره گرفت. مقدار $\rho_s(1)$ نشان می‌دهد که روش s برای چند درصد مسایل در کمترین زمان (تعداد محاسبات) ممکن جواب می‌دهد. لذا، اگر منظور از بهترین روش، روشی باشد که در کمترین زمان (تعداد محاسبات) برای بیشتر مسایل جواب دهد، از بین روش‌های موجود روشی را انتخاب می‌کنیم که $\rho_s(1)$ بزرگتری داشته باشد. $\rho_s(1)$ متناظر با نقطه شروع روی محور عمودی است.

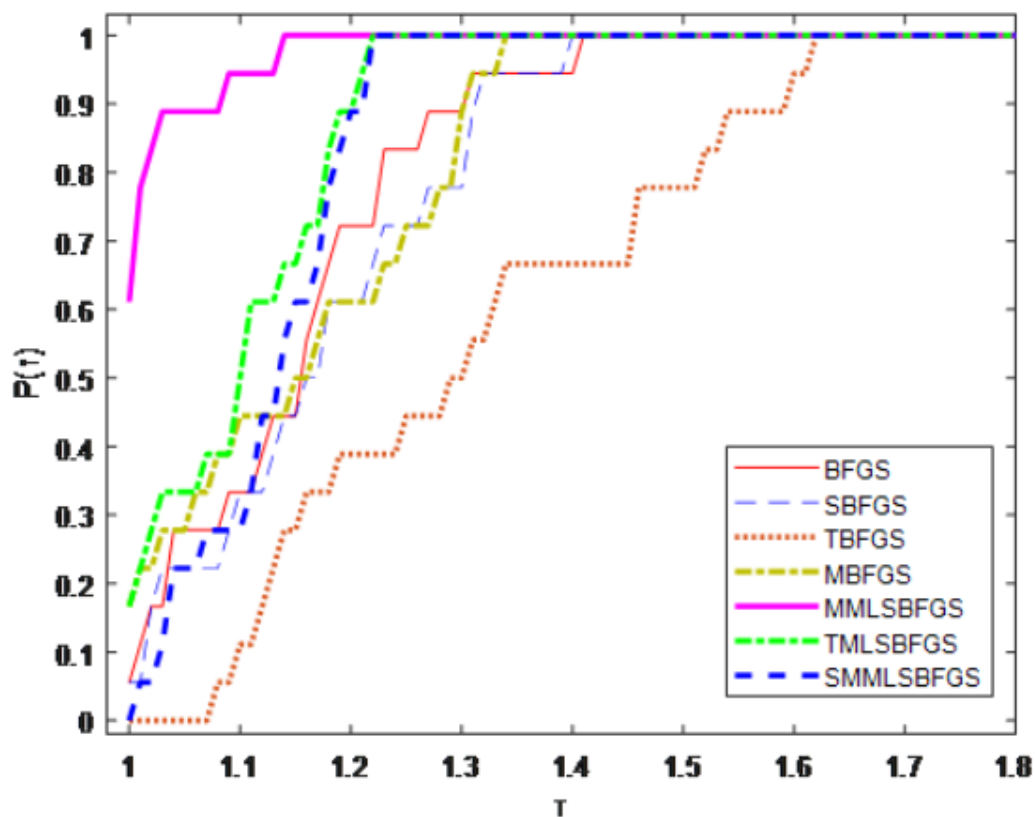
تعریف $\rho_s(\tau)$ برای مقادیر بزرگ τ به کمی دقت نیاز دارد. می‌دانیم که $r_{p,s} \in [1, r_M]$ و $r_{p,s} = r_M$ اگر و تنها اگر روش s نتواند مساله p را حل کند. لذا می‌توان گفت $\rho_s(r_M) = 1$ و

$$\rho_s^* = \lim_{\tau \rightarrow r_M^-} \rho_s(\tau).$$

^۵Dolan and Moré

ρ_s^* نشان می‌دهد که روش s در مدت زمان مجاز، حداکثر چند درصد مسایل را حل می‌کند. لذا، اگر هدف انتخاب روشی باشد که برای بیشتر مسایل جواب می‌دهد، روشی را انتخاب می‌کنیم که ρ_s^* بزرگتری داشته باشد. متناظر با این مقدار، نمودار ρ_s به خط راست افقی تبدیل می‌شود.

با توجه به توضیحات داده شده، نمودار دولان – موره را بر حسب زمان اجرای الگوریتم برای مقایسه کارایی الگوریتم‌های مورد آزمون رسم می‌کنیم. در شکل ۲.۴ مشاهده می‌شود که نمودار کارایی الگوریتم‌های MMLS BFGS و TMLS BFGS بالاتر از نمودار کارایی الگوریتم‌های دیگر است. پس، نتیجه می‌گیریم که الگوریتم‌های MMLS BFGS و TMLS BFGS در مقایسه با الگوریتم‌های دیگر کارایی بهتری دارند.



شکل ۲.۴: نمودار دولان – موره بر حسب زمان اجرا برای مقایسه کارایی الگوریتم‌های مورد آزمون

فصل ۵

حل مساله مکان‌یابی توافقی در مقیاس بزرگ با رویکرد تصادفی

۱.۵ مقدمه

در این فصل یک مساله مکان‌یابی توافقی مقیاس بزرگ تحت نرم ℓ_p را در نظر می‌گیریم. حل مسایل مقیاس بزرگ با استفاده از روش‌های بهینه‌سازی قطعی به دلیل سرعت همگرایی کند نامعقول است زیرا محاسبه گرادیان و معکوس هسی یا تقریبی از معکوس هسی هزینه محاسباتی بالایی را بر الگوریتم تحمیل می‌کند به همین دلیل روش‌های بهینه‌سازی تصادفی در مسایل مقیاس بزرگ جایگاه ویژه‌ای به خود اختصاص داده‌اند. یکی از روش‌های بهینه‌سازی تصادفی مرتبه دوم، نسخه تصادفی روش LBFGS است که در عمل نتایج مطلوبی برای مسایل مقیاس بزرگ به دست می‌دهد. بنابراین این روش را برای حل مساله مکان‌یابی توافقی به کار می‌گیریم و نتایج عددی حاصل از آن را ارائه می‌دهیم.

فرض کنید n نقطه موجود باشند و $A_1, \dots, A_n$ نیز شعاع‌های توافقی متناظر با نقاط تقاضای $A_1, \dots, A_n$ باشند. همچنین فرض کنید w_i وزن مربوط به هر نقطه تقاضای A_i باشد. بنابراین یک مساله مکان‌یابی وبر توافقی تحت تابع زیان مربعات خطا و با نرم ℓ_p ، به

صورت زیر مدل‌بندی می‌شود:

$$\min f(X) = \frac{1}{n} \sum_{i=1}^n w_i \cdot (d(X, A_i) - R_i)^2 + U(X) \quad (۱.۵)$$

به طوری‌که:

$$d(X, A_i) = \left(\left((x - a_i)^2 + \epsilon \right)^{\frac{p}{2}} + \left((y - b_i)^2 + \epsilon \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \quad (۲.۵)$$

و $U : \mathbb{R}^d \rightarrow \mathbb{R}$ بخش منظم‌ساز^۱ است که تابعی محدب است و در این پژوهش منظم‌ساز ℓ_2 در نظر گرفته می‌شود.

در مسایل مقیاس بزرگ n عددی بزرگ است، به همین علت محاسبه هسی (یا تقریب هسی) و گرادیان برای این مسایل دارای هزینه محاسباتی بالایی است که می‌توان با استفاده از روش‌های بهینه‌سازی تصادفی این هزینه را کاهش داد. اکنون، با قرار دادن

$$f_i = w_i \cdot (d(X, A_i) - R_i)^2 \quad (۳.۵)$$

برای زیرمجموعه $S \subseteq \{1, \dots, n\}$ تابع f_S به صورت

$$f_S(X) = \frac{1}{|S|} \sum_{i \in S} f_i(X) + U(X) \quad (۴.۵)$$

تعریف می‌شود. توجه شود که در ادامه بخش منظم‌ساز را درون تابع اصلی در نظر گرفته و آن را به صورت جداگانه در نظر نمی‌گیریم. یعنی، مسئله (۱.۵) به صورت $\min f(X) = \frac{1}{n} \sum_{i=1}^n f_i(X)$ فرض می‌شود.

به منظور شرح قضایا و نتایج نظری، لازم است ابتدا تابع λ - محدب قوی^۲ و Λ - هموار^۳ را به صورت زیر تعریف کنیم.

تعریف ۱.۱.۵ [۵] تابع محدب $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ، λ - محدب قوی و Λ - هموار است هرگاه برای هر $X, Z \in \mathbb{R}^d$ نابرابری‌های زیر برقرار باشند:

$$\nabla f(X)^T (Z - X) + \frac{\Lambda}{2} \|Z - X\|^2 \geq f(Z) - f(X) \geq \nabla f(X)^T (Z - X) + \frac{\lambda}{2} \|Z - X\|^2.$$

در بخش بعدی، ایده‌های اصلی تکنیک‌های بهینه‌سازی تصادفی برای مسایل مقیاس بزرگ خلاصه خواهد شد. در بخش سوم، مباحث نظری مرتبط به یکی از این ایده‌ها را به صورت مبسوط ارائه می‌دهیم و به توضیح نسخه تصادفی روش LBFGS می‌پردازیم و در بخش آخر نتایج عددی حاصله را برای مساله مکان‌یابی توافقی (۱.۵) ارائه می‌دهیم.

^۱regularizer term

^۲ λ -strongly convex

^۳ Λ -smooth

جدول ۱.۵: خلاصه‌ای از روش‌های بهینه‌سازی تصادفی	
روش‌های مرتبه اول (۱.۲.۵)	روش گرادیان کاهشی تصادفی روش‌های کاهش واریانس (SVRG, SAG)
روش‌های مرتبه دوم (۲.۲.۵)	روش‌های نیوتن بدون استفاده از ماتریس هسی روش‌های شبه نیوتن (LBFGS)
روش نیوتن منظم‌سازی شده مکعبی (۳.۲.۵)	
روش گرادیان کاهشی مکعبی با سرعت خطی روش منظم‌سازی شده مکعبی با واریانس کاهش یافته	
برخی از روش‌های متداول دیگر (۴.۲.۵)	
روش‌های گرادیان با گشتاور روش‌های گرادیان تسریع شده روش‌های مختصات کاهشی	

۲.۵ روش‌های بهینه‌سازی تصادفی

این روش‌ها به سه دسته کلی تقسیم می‌شوند: روش‌های مرتبه اول^۴، روش‌های مرتبه دوم^۵ و روش‌های نیوتن منظم‌سازی شده مکعبی^۶. خلاصه‌ای از این تکنیک‌ها در جدول ۱.۵ ارائه شده است.

۱.۲.۵ روش‌های مرتبه اول

روش‌های مرتبه اول در الگوریتم‌های بهینه‌سازی مورد توجه هستند زیرا می‌توان آن‌ها را در زمان متناسب با بعد مساله پیاده‌سازی کرد [۶۰]. در واقع، روش‌های مرتبه اول براساس گرادیان (یعنی، اطلاعات مرتبه اول تابع) عمل می‌کنند.

۱.۱.۲.۵ روش گرادیان کاهشی تصادفی (SGD)

روش گرادیان کاهشی^۷ یک روش تکراری مرتبه اول است که در هر تکرار با حرکت در خلاف جهت گرادیان تابع، سعی در کاهش مقدار تابع دارد و به صورت خطی به نقطه بهینه همگرا می‌شود:

$$X_t = X_{t-1} - \eta \sum_{i=1}^n \nabla f_i(X_{t-1})$$

اگرچه ممکن است روش‌های سنتی مبتنی بر گرادیان برای حل مسایل در مقیاس کوچک (n کوچک) مؤثر باشند، اما برای مسایل مقیاس بزرگ هزینه محاسباتی بالایی ایجاد می‌شود زیرا

^۴ first-order methods

^۵ second-order methods

^۶ cubic regularized Newton methods

^۷ Gradient descent method

در هر تکرار باید n گرادیان محاسبه شود. پس منطقی است که از نوع تصادفی الگوریتم‌های مبتنی بر گرادیان استفاده شود.

این روش توسط رابینز^۸ و مونرو^۹ [۱۱۴] با نام روش گرادیان کاهشی تصادفی^{۱۰} در سال ۱۹۵۱ ارائه شده است. در این روش در هر تکرار عدد i را به صورت تصادفی از میان اعداد $\{1, 2, \dots, n\}$ انتخاب می‌کنیم و پارامتر مساله به صورت تکراری طبق قاعده زیر بهنگام می‌شود

$$X_t = X_{t-1} - \eta \nabla f_i(X_{t-1})$$

که در آن η طول گام (نرخ یادگیری^{۱۱}) نامیده می‌شود. مزیت این روش این است که در هر تکرار فقط یک گرادیان ∇f_i محاسبه می‌شود که هزینه محاسباتی آن $\frac{1}{n}$ هزینه محاسباتی روش گرادیان کاهشی استاندارد (FGD)^{۱۲} است. اما استفاده از گرادیان مرتبط به یک داده به عنوان برآوردگری نااریب به جای کل گرادیان باعث ایجاد واریانس می‌شود که سرعت همگرایی را کاهش می‌دهد. روش SGD حتی برای توابع محدب قوی نیز به صورت زیرخطی همگرا می‌شود [۶۱].

۲.۱.۲.۵ روش‌های کاهش واریانس

اخیراً دستاورد مهمی در روش‌های مرتبه اول به دست آمده است که با کاهش واریانس، زمان اجرای الگوریتم را بهبود می‌بخشد. این روش‌ها برای توابع محدب قوی به صورت خطی همگرا می‌شوند که موجب بهبود سرعت زیرخطی SGD می‌شود. این دستاورد با افزایش هزینه محاسباتی یا افزایش حافظه مورد نیاز برای ذخیره‌سازی حاصل می‌شود. این الگوریتم‌ها شامل گرادیان تصادفی میانگین (SAG)^{۱۳} [۱۱۶، ۳۹، ۱۱۱] و گرادیان تصادفی با واریانس کاهش یافته (SVRG)^{۱۴} [۲۳، ۱۲۳، ۷۶] هستند.

گرادیان تصادفی میانگین در این روش می‌توان به کمک کاهش واریانس برآوردگر، ضمن حفظ هزینه محاسباتی تکرار روش SGD به نرخ همگرایی خطی دست پیدا کرد. این روش، ترکیبی از هزینه تکرار پایین روش SGD و نرخ همگرایی خطی مشابه روش گرادیان کاهشی استاندارد است و از تکرارهایی به شکل

$$X_t = X_{t-1} - \frac{\eta}{n} \sum_{i=1}^n y_i^t$$

^۸Robbins

^۹Monro

^{۱۰}Stochastic Gradient Descent method (SGD)

^{۱۱}learning rate

^{۱۲}Full Gradient Descent (FGD)

^{۱۳}Stochastic Average Gradient

^{۱۴}Stochastic Variance Reduced Gradient

استفاده می‌کند که در هر تکرار داده i_t -ام به صورت تصادفی انتخاب می‌شود و قرار می‌دهد:

$$y_i^t = \begin{cases} \nabla f_i(X_t) & i = i_t \text{ اگر} \\ y_i^{t-1} & o.w. \end{cases}$$

یعنی، در هر تکرار فقط گرادیان مربوط به داده i -ام محاسبه می‌شود و به جای بقیه گرادیان‌ها، همان مقادیر تکرار قبل در نظر گرفته می‌شود. در واقع، در این روش مانند روش FGD در هر تکرار گرادیان مربوط به همه داده‌ها به کار گرفته می‌شود. اما، مانند روش SGD در هر تکرار فقط گرادیان مربوط به یک داده آموزشی محاسبه می‌شود و هزینه تکرارها مستقل از n است. روش SAG علاوه بر هزینه کم محاسباتی مانند روش FGD نرخ همگرایی خطی دارد. این بدان معناست که این تکرار با دسترسی به داده i_t -ام و نگهداری مقادیر گرادیان محاسبه شده اخیر برای هر داده آموزشی و اختصاص حافظه‌ای برای این مقادیر، سرعت همگرایی سریع‌تری را از روش‌های استاندارد SGD به دست می‌آورد.

گرادیان تصادفی با واریانس کاهش یافته این روش به صورت چرخه‌ای (دوره‌ای) عمل می‌کند، هر چرخه با یک گرادیان کامل در X_t شروع می‌شود. یعنی، $\nabla f(X_t) = \frac{1}{n} \sum_{i=1}^n \nabla f_i(X_t)$. سپس برای حلقه درونی ابتدا قرار می‌دهد $\tilde{X}_1 = X_t$ و تکرارهای حلقه درونی به صورت $\tilde{X}_{k+1} = \tilde{X}_k - \eta \tilde{g}_k$ بهنگام می‌شوند که در آن

$$\tilde{g}_k = \nabla f_{i_k}(\tilde{X}_k) - (\nabla f_{i_k}(X_t) - \nabla f(X_t)), \quad (5.5)$$

و i_k به صورت تصادفی از $\{1, \dots, n\}$ انتخاب می‌شود. می‌توان معادله (5.5) را به صورت زیر تفسیر نمود. بدیهی است که امید ریاضی $\nabla f_{i_k}(X_t)$ برابر است با $\nabla f(X_t)$ که در آن متغیر تصادفی $i_k \in \{1, \dots, n\}$ به صورت تصادفی انتخاب می‌شود. بنابراین، $\nabla f_{i_k}(X_t) - \nabla f(X_t)$ میزان اریبی در برآورد $\nabla f_{i_k}(X_t)$ است. پس، الگوریتم به طور تصادفی $i_k \in \{1, \dots, n\}$ را انتخاب کرده و گرادیان تصادفی $\nabla f_{i_k}(\tilde{X}_k)$ را در هر تکرار محاسبه کرده و آن را طبق اریبی متناظر تصحیح می‌کند. توجه داشته باشید که $\tilde{X}_k$ نقطه فعلی در حلقه درونی است. با این تفاسیر، می‌توان گفت که $\tilde{g}_k$ برآوردگر نااریب $\nabla f(\tilde{X}_k)$ است. توجه داشته باشید که اگر مانند SGD، $\tilde{g}_k = \nabla f_{i_k}(\tilde{X}_k)$ انتخاب شود، واریانس بیشتری نسبت به این روش ایجاد می‌شود. بنابراین، یک تکرار SVRG بسیار پرهزینه‌تر از یک تکرار SGD است و در واقع با یک تکرار FGD قابل مقایسه است. نمونه‌ای از این الگوریتم را می‌توان در الگوریتم ۷ مشاهده کرد.

۲.۲.۵ روش‌های مرتبه دوم

روش‌های مرتبه دوم، مانند روش نیوتن، به دلیل نرخ همگرایی درجه دوم بسیار رایج هستند، اما محاسبه ماتریس هسی (به صورت کامل) و وارون آن به ویژه برای مسایل مقیاس بزرگ

الگوریتم ۷ الگوریتم SVRG [۲۳]

- ۱: ورودی: $X_1, f(X) = \frac{1}{n} \sum_{i=1}^n f_i(X), T, m$
- ۲: برای $t = 1$ تا T
- ۳: گرادیان کامل $\nabla f(X_t)$ را محاسبه کن.
- ۴: قرار ده $\tilde{X}_1 = X_t$
- ۵: برای $k = 1$ تا m
- ۶: i_k را به صورت تصادفی از $\{1, \dots, n\}$ انتخاب کن.
- ۷: قرار ده $\tilde{g}_k = \nabla f_{i_k}(\tilde{X}_k) - (\nabla f_{i_k}(X_t) - \nabla f(X_t))$
- ۸: قرار ده $\tilde{X}_{k+1} = \tilde{X}_k - \eta \tilde{g}_k$
- ۹: پایان حلقه
- ۱۰: انتخاب الف: $X_{t+1} = \tilde{X}_{m+1}$
- ۱۱: انتخاب ب: $X_{t+1} = \frac{1}{m} \sum_{k=1}^m \tilde{X}_{k+1}$
- ۱۲: انتخاب ج: k را به صورت تصادفی و با توزیع یکنواخت از $\{1, \dots, m\}$ انتخاب کن و قرار ده $X_{t+1} = \tilde{X}_{k+1}$
- ۱۳: پایان حلقه

بسیار پرهزینه است (به ترتیب $O(nd^2)$ و $O(d^3)$). اخیراً، الگوریتم نیوسمپ^{۱۵} [۴۵] ایده‌هایی را برای حل این دو مشکل پیشنهاد داده است. این الگوریتم برای تقریب هسی از تکنیک‌های نمونه‌برداری تصادفی استفاده می‌کند و برای کاهش هزینه محاسبه وارون هسی از تکنیک‌های تصویرسازی رتبه پایین^{۱۶} بهره می‌برد.

۱.۲.۲.۵ روش‌های نیوتن بدون استفاده از ماتریس هسی

یافتن جهت جستجوی s_t در روش نیوتن مستلزم حل معادله نیوتن، یعنی، $H_t s_t = -g_t$ ، است که دارای هزینه محاسباتی و نیازمند حافظه برای ذخیره‌سازی است. بنابراین، به جای حل دقیق معادله نیوتن، می‌توان از روش‌های غیردقیق مانند روش گرادیان مزدوج نیوتن (NCG)^{۱۷} برای حل این معادله و به دست آوردن نرخ همگرایی زبرخطی^{۱۸} استفاده کرد. در روش گرادیان مزدوج نیوتن، نیازی به دسترسی دقیق به ماتریس هسی نیست، بلکه حاصل ضرب بردار- هسی کافی است که همین موضوع دلیل نام‌گذاری این روش است. برای مطالعات بیشتر می‌توان به مراجع [۹۰] و [۲۷] مراجعه کرد.

^{۱۵}NewSamp

^{۱۶}low-rank projection

^{۱۷}Newton Conjugate Gradient (NCG) method

^{۱۸}superlinear

هم‌چنین، در سال ۲۰۱۷، لیزا^{۱۹} [۵] از روشی خاص برای محاسبه معکوس هسی استفاده کرد و یک برآوردگر ناریب برای معکوس هسی به دست آورد که این روش دارای سرعت همگرایی خطی است.

۲.۲.۲.۵ روش‌های شبه نیوتن

همان‌طور که قبلاً نیز اشاره شد، محاسبه ماتریس هسی پرهزینه است. روش‌های دیگری که برای کاهش این هزینه محاسباتی استفاده می‌شود، روش‌های شبه نیوتن هستند که سرشته آن‌ها الگوریتم BFGS (برویدن^{۲۰}، فلچر^{۲۱}، گلدفارب^{۲۲} و شانو^{۲۳}) و نوع حافظه محدود^{۲۴} آن (L-BFGS) است [۱۰۵]. در این روش، تقریب هسی با استفاده از تغییرات گرادیان محاسبه می‌شود. در حقیقت، ثابت شده است که نوع حافظه محدود این روش برای مسایل با میلیون‌ها متغیر، کارا است.

برای مسایل مقیاس بزرگ، از تکنیک‌های تصادفی در این الگوریتم‌ها استفاده شده است و اشکال مختلف روش‌های تصادفی نیوتن ارائه شده است. در سال ۲۰۰۷، برای اولین بار، یک مدل تصادفی ساده برای الگوریتم BFGS با استفاده از روش نمونه‌گیری تصادفی برای محاسبه تغییرات گرادیان ارائه شد [۱۱۹]. علاوه بر این، مختاری و ریبیرو^{۲۵} یک الگوریتم تصادفی BFGS منظم‌سازی شده ارائه کرده است [۹۵]. هم‌چنین، برد و همکاران [۲۸] برآوردگرهایی برای میانگین اطلاعات انحنا با استفاده از تکنیک حاصل ضرب هسی-بردار محاسبه کرده است. اما همه این الگوریتم‌ها نرخ همگرایی زیرخطی دارند. سرانجام، در سال ۲۰۱۶، دستاورد مهمی توسط موریتز^{۲۶} و همکاران حاصل شد. آن‌ها با ایده‌های کاهش واریانس (شبه به روش SVRG) و تکنیک دسته‌ای کوچک^{۲۷} به سرعت همگرایی خطی دست یافتند [۹۸] که در این پایان‌نامه به شرح این روش خواهیم پرداخت و نتایج عددی حاصل را ارائه خواهیم داد.

۳.۲.۵ روش‌های نیوتن منظم‌سازی شده مکعبی

یکی دیگر از روش‌های متداول بهینه‌سازی، روش نیوتن منظم‌سازی شده مکعبی است که توسط نسترو^{۲۸} و پلیاک^{۲۹} [۱۰۲] معرفی شده است. هدف این روش، کمینه‌سازی تابع

^{۱۹}LissA

^{۲۰}Broyden

^{۲۱}Fletcher

^{۲۲}Goldfarb

^{۲۳}Shanno

^{۲۴}limited memory

^{۲۵}Mokhtari and Ribeiro

^{۲۶}Moritz

^{۲۷}mini-batch

^{۲۸}Nesterov

^{۲۹}Polyak

مکعبی

$$m(h, X) = \nabla f(X)^T h + \frac{1}{2} h^T \nabla^2 f(X) h + \frac{L}{6} \|h\|^3 \quad (۶.۵)$$

است که در آن $L > 0$ پارامتر منظم‌سازی است. بنابراین، تکرارها به صورت زیر به‌نگام می‌شوند

$$X_{t+1} = X_t + h(X_t),$$

که $h(X) = \arg \min_{h \in \mathbb{R}^d} m(h, X)$. بدین منظور تکنیک‌های مختلفی ارائه شده است که در این بخش دو تکنیک جالب شرح داده می‌شود. برای مطالعهٔ بیش‌تر می‌توان به [۱۴۱، ۱۳۱، ۱۴۲] مراجعه کرد.

در سال ۲۰۱۷، آگاروال^{۳۰} و همکاران [۴] به صورت تقریبی مقادیر ویژه و بردار ویژه ماتریس هسی را با استفاده از روش توانی^{۳۱} و آنالیز مؤلفه‌های اصلی^{۳۲} محاسبه کردند. در این روش از تکنیک‌های مفیدی برای تقریب معکوس ماتریس استفاده می‌شود و به کمک این تقریب یک جهت جستجوی کاهشی برای کمینه‌سازی تابع مکعبی (۶.۵) به دست آوردند. لازم به ذکر است که این روش برای توابع نامحدب پیشنهاد شده است و دارای سرعت همگرایی خطی است.

در سال ۲۰۱۸، ژو^{۳۳} و همکاران [۱۴۶] برآوردگرهای جالبی برای گرادیان و هسی با استفاده از تکنیک‌های نمونه‌برداری تصادفی و کاهش واریانس ارائه داده‌اند. کمینه‌سازی تابع مکعبی با کمک این برآوردگرها انجام شده است. در واقع در این مقاله از ایده روش SVRG برای کاهش واریانس برآوردگر گرادیان و هسی استفاده شده است اما این روش با روش SVRG دو تفاوت مهم دارد. تفاوت اول در روش نمونه‌برداری دسته‌ای کوچک است که برای محاسبه برآوردگر از بیش از یک گرادیان و هسی تصادفی استفاده می‌کند و تفاوت دوم در این است که از اطلاعات مرتبهٔ دوم نیز برای محاسبهٔ جهت جستجوی کاهشی برای کمینه‌سازی تابع مکعبی استفاده شده است.

۴.۲.۵ برخی از روش‌های متداول دیگر

روش‌های بهینه‌سازی دیگری نیز وجود دارند که دستاوردهای مفیدی کسب کرده‌اند. این زیربخش به معرفی این روش‌ها می‌پردازد که شامل روش‌های گرادیان با گشتاور^{۳۴}، روش‌های گرادیان تسریع‌شده^{۳۵} و روش‌های مختصات کاهشی^{۳۶} هستند.

^{۳۰} Agarwal

^{۳۱} power method

^{۳۲} Principal Component Analysis (PCA)

^{۳۳} Zhou

^{۳۴} gradient methods with momentum

^{۳۵} accelerated gradient methods

^{۳۶} coordinate descent methods

۱.۴.۲.۵ روش‌های گرادیان با گشتاور

در روش‌های گرادیان با گشتاور هر گام به صورت ترکیبی از جهت تندترین کاهش و جهت تکرار اخیر انتخاب می‌شود. در ابتدا، یک نقطه آغازین X_0 و دنباله‌ای از اسکالرهایی $\{\alpha_t\}$ و $\{\beta_t\}$ در نظر گرفته می‌شوند که از پیش تعیین شده یا به صورت پویا^{۳۷} انتخاب می‌شوند. سپس تکرارها به صورت زیر بهنگام می‌شوند

$$X_1 = X_0, \quad X_{t+1} = X_t - \alpha_t \nabla f(X_t) + \beta_t (X_t - X_{t-1}). \quad (۷.۵)$$

سمت راست معادله (۷.۵) گشتاور نامیده می‌شود که بطور بازگشتی حرکت الگوریتم را در مسیرهای جستجوی قبلی حفظ می‌کند که از متداول‌ترین روش‌های گرادیان با گشتاور می‌توان به روش گرادیان مزدوج اشاره کرد [۱۰۵]. برای مسایل با مقیاس بزرگ، به جای گرادیان کامل در (۷.۵) از گرادیان تصادفی استفاده می‌شود.

الگوریتم کاتیوشا^{۳۸} [۶، ۸۲] با استفاده از ایده گشتاور نسترو [۱۰۳]، مفهوم گشتاور منفی^{۳۹} را معرفی کرد که نوعی روش کاهش واریانس است. در حقیقت، کاتیوشا الگوریتم SVRG را برای مسایل محدب قوی با استفاده از تکنیک دسته‌ای کوچک تسریع می‌کند.

۲.۴.۲.۵ روش‌های گرادیان تسریع‌شده

فرمول این روش مشابه (۷.۵) است و ایده تسریع شده آن توسط نسترو [۱۰۴] ارائه شده است. هر تکرار دارای شکل زیر است:

$$\begin{aligned} \tilde{X}_t &= X_t + \beta_t (X_t - X_{t-1}), \\ X_{t+1} &= \tilde{X}_t - \alpha_t \nabla f(\tilde{X}_t), \end{aligned}$$

که می‌توان به سادگی به فرمول زیر دست یافت:

$$X_{t+1} = X_t - \alpha_t \nabla f(X_t + \beta_t (X_t - X_{t-1})) + \beta_t (X_t - X_{t-1}). \quad (۸.۵)$$

فرمول‌های (۷.۵) و (۸.۵) را می‌توان بدین صورت تفسیر کرد. در (۷.۵) ابتدا گام تندترین کاهش در نظر گرفته می‌شود و سپس بخش گشتاور اعمال می‌شود، در حالی که در (۸.۵) ابتدا بخش گشتاور در نظر گرفته می‌شود، سپس گام تندترین کاهش اعمال می‌شود (با گرادیان محاسبه شده در $\tilde{X}_t$ ، نه در X_t).

برای مسایل با مقیاس بزرگ، گرادیان تصادفی به جای گرادیان کامل در (۸.۵) استفاده می‌شود. این تکنیک در دسته وسیعی از الگوریتم‌ها به کار گرفته شده است. یکی از این

^{۳۷}dynamical

^{۳۸}Katyusha

^{۳۹}negative momentum

الگوریتم‌ها، کاتالیست^{۴۰} [۸۶] است که از این استراتژی برای تسریع روش گرادیان استفاده می‌کند.

۳.۴.۲.۵ روش‌های مختصات کاهش‌ی

روشی دیگر، که از قدیمی‌ترین روش‌های بهینه‌سازی است، روش مختصات کاهش‌ی نام دارد. همان‌طور که از نام آن پیدا است، این روش با در نظر گرفتن گام‌هایی در راستای جهت‌های مختصات عمل می‌کند. این روش سعی می‌کند تابع هدف را برحسب یکی از متغیرها کمینه کند در حالی که بقیه متغیرها ثابت نگه داشته می‌شوند، سپس سایر متغیرها به طور مشابه و به صورت تکراری در گام‌های بعد بهنگام می‌شوند. تکرارهای این روش به شکل زیر هستند:

$$X_{t+1} = X_t - \alpha_t \nabla_{i_t} f(X_t) e_{i_t}, \quad \nabla_{i_t} f(X_t) := \frac{\partial f}{\partial X^{i_t}}(X_t). \quad (9.5)$$

توجه داشته باشید که X_{i_t} مؤلفه i_t - ام بردار X را نشان می‌دهد و e_{i_t} نشانگر بردار مختصات i_t - ام است برای $i_t \in \{1, \dots, d\}$.

برای مسایل با مقیاس بزرگ، نسخه تصادفی روش مختصات کاهش‌ی مورد بحث قرار می‌گیرد که روش مختصات افزایشی دوگان تصادفی (SDCA)^{۴۱} نامیده می‌شود [۱۲۱]. برای این منظور، جایگزین کردن $\nabla_{i_t} f(X_t)$ در (۹.۵) با نسخه تصادفی آن متداول نیست، زیرا می‌توان به سادگی یک گرادیان تصادفی d بعدی را برای اعمال روش SGD محاسبه کرد. بنابراین، می‌توان با استفاده از بیشینه‌سازی مسئله دوگان آن و استفاده از گام‌های مثبت گرادیان به جای منفی گرادیان، به نسخه تصادفی این روش دست یافت. SDCA دارای نرخ همگرایی خطی با ثابت وابسته به بعد مساله، یعنی d ، است. به علاوه، برای مطالعه نسخه تسریع‌شده روش مختصات افزایشی دوگان تصادفی می‌توانید [۱۲۲] و [۸۷] را ببینید.

۳.۵ نسخه تصادفی الگوریتم LBFGS

در این فصل با کمک یک روش تصادفی مرتبه دوم به کمینه‌سازی خطای مساله مکان‌یابی می‌پردازیم. این روش در حالاتی که مجموعه داده دارای اندازه بزرگی هست جواب‌های قابل قبولی در زمان کمتری به دست می‌دهد. بدین منظور به معرفی نسخه تصادفی روش LBFGS می‌پردازیم و خطای مساله مکان‌یابی را با استفاده از این روش کمینه می‌کنیم و نتایج حاصل را ارائه می‌دهیم.

حل مسایل مقیاس بزرگ با استفاده از روش‌های بهینه‌سازی قطعی به دلیل سرعت همگرایی کند نامعقول است زیرا محاسبه گرادیان و معکوس هسی یا تقریبی از معکوس هسی هزینه محاسباتی بالایی را متحمل الگوریتم می‌کند به همین دلیل روش‌های بهینه‌سازی تصادفی در

^{۴۰}Catalyst

^{۴۱}Stochastic Dual Coordinate Ascent (SDCA)

مسایل مقیاس بزرگ جایگاه ویژه‌ای به خود اختصاص داده‌اند. یکی از روش‌های بهینه‌سازی تصادفی مرتبه دوم، نسخه تصادفی روش LBFGS است که در عمل نتایج مطلوبی برای مسایل مقیاس بزرگ به دست می‌دهد.

اگرچه اغلب الگوریتم‌های تصادفی سرعت پردازش و اجرای بالایی دارند اما واریانس حاصل از برآورد گرادیان موجب کاهش همگرایی در نزدیکی جواب بهینه می‌شود. به عنوان مثال در روش SGD حتی اگر نقطه شروع را در نزدیکی نقطه بهینه انتخاب کنیم، ممکن است به تابع هدف با مقدار بدتری دست یابیم. به همین دلیل برای تضمین همگرایی نیاز به طول گام‌های نزولی (کم‌شونده) داریم. یکی از راه‌های حل این مشکل، استفاده از روش‌هایی است که موجب کاهش واریانس برآوردگر گرادیان می‌شود و سرعت الگوریتم را افزایش می‌دهد.

حال به شرح نسخه تصادفی روش LBFGS می‌پردازیم و از یک تکنیک مشابه روش SVRG برای کاهش واریانس برآوردگر گرادیان استفاده می‌کنیم. این الگوریتم در مسایل مقیاس بزرگ عملکرد مطلوبی دارد و به دلیل استفاده از ایده تکنیک کاهش واریانس دارای سرعت همگرایی بالایی برای توابع محدب قوی است.

۱.۳.۵ الگوریتم SLBFGS

در این قسمت به شرح جزئیات نسخه تصادفی الگوریتم LBFGS می‌پردازیم و نکات مهم آن را بیان می‌کنیم.

تکرارهای این الگوریتم تصادفی از برآوردگرهایی تصادفی به جای گرادیان کامل و هسی کامل تابع اصلی استفاده می‌کنند که این برآوردگرها را به ترتیب با نمادهای ∇f_S و $\nabla^2 f_T$ نمایش می‌دهیم که در آن $S, T \subseteq \{1, \dots, n\}$ دو نمونه تصادفی متفاوت هستند. حال قرار می‌دهیم $|T| = b_H$ و $|S| = b$.

مشابه روش SVRG، به منظور کاهش واریانس برآوردگر تصادفی گرادیان در هر تکرار بیرونی یک بار گرادیان کامل محاسبه می‌شود و تکرارها طبق قاعده

$$X_{k+1} = X_k - \eta_k H_k v_k$$

به‌هنگام می‌شوند که در آن H_k نشان‌دهنده تقریبی از معکوس هسی و v_k بیانگر برآوردگر تصادفی گرادیان با واریانس کاهش یافته در تکرار k -ام است.

تکرارهای این الگوریتم تصادفی در الگوریتم ۸ نشان داده شده است. این الگوریتم دارای چند پارامتر خاص هست که در ادامه به توصیف این پارامترها می‌پردازیم. پارامتر η بیانگر طول گام الگوریتم است. عدد صحیح مثبت m که تعداد تکرار حلقه درونی را مشخص می‌کند نشان‌دهنده تعداد محاسبه گرادیان کامل است که همان‌طور که قبلاً نیز بیان کردیم، استفاده از گرادیان کامل موجب کاهش واریانس برآوردگر تصادفی گرادیان می‌شود. به‌علاوه در هر L مرتبه تکرار الگوریتم، تقریبی جدید برای معکوس هسی محاسبه می‌شود. بردار s_r میانگین

جهاتی که در طول $2L$ تکرار اخیر الگوریتم به دست آمده است را ذخیره می‌کند. بردار y_r از حاصل ضرب بردار s_r در برآوردگر تصادفی ماتریس هسی حاصل می‌شود. توجه می‌کنیم که یکی از تفاوت های الگوریتم LBFGS و SLBFGS در چگونگی تعریف بردار y_r است. در الگوریتم LBFGS بردار y_r به صورت تفاضل گرادیان‌ها تعریف می‌شود اما در نسخه تصادفی تعریف ذکر شده عملکرد بهتری ایجاد می‌کند. حال برای محاسبه تقریب‌هایی برای معکوس هسی از بردارهای y_r و s_r استفاده می‌کنیم. بدین منظور، قرار می‌دهیم $\rho_j = 1/s_j^T y_j$ و $M' = \min\{M, r\}$. به صورت بازگشتی تقریب‌هایی برای محاسبه معکوس هسی به صورت

$$H_r^{(j)} = (I - \rho_j s_j y_j^T)^T H_r^{(j-1)} (I - \rho_j s_j y_j^T) + \rho_j s_j s_j^T, \quad j \in \{r - M' + 1, \dots, r\} \quad (10.5)$$

تعریف می‌کنیم. با $H_r^{(r-M')} = \frac{s_r^T y_r}{\|y_r\|^2}$ شروع می‌کنیم و قرار می‌دهیم $H_r = H_r^{(r)}$. توجه می‌کنیم که قاعده فوق معین مثبت بودن را حفظ می‌کند ($\rho_j > 0$).

الگوریتم ۸ الگوریتم SLBFGS [۹۸]

- ۱: ورودی: X_0, m, L, η
- ۲: قرار ده $r = 0$
- ۳: قرار ده $H_0 = I$
- ۴: برای $k = 0$ تا $\dots$
- ۵: گرادیان کامل $\mu_k = \nabla f(X_k)$ را محاسبه کن.
- ۶: قرار ده $x_0 = X_k$
- ۷: برای $t = 0$ تا $m - 1$
- ۸: $S_{k,t} \subseteq \{1, \dots, n\}$ را به صورت تصادفی انتخاب کن.
- ۹: گرادیان تصادفی $\nabla f_{S_{k,t}}(x_t)$ را محاسبه کن.
- ۱۰: قرار ده $v_t = \nabla f_{S_{k,t}}(x_t) - \nabla f_{S_{k,t}}(X_k) + \mu_k$
- ۱۱: قرار ده $x_{t+1} = x_t - \eta H_r v_t$
- ۱۲: اگر $t \equiv 0$ به پیمانه L آن گاه
- ۱۳: $r = r + 1$
- ۱۴: قرار ده $u_r = \frac{1}{L} \sum_{j=t-L}^{t-1} x_j$
- ۱۵: $T_r \subseteq \{1, \dots, n\}$ را برای محاسبه تقریب $\nabla^2 f_{T_r}(u_r)$ انتخاب کن.
- ۱۶: $s_r = u_r - u_{r-1}$ را محاسبه کن.
- ۱۷: $y_r = \nabla^2 f_{T_r}(u_r) s_r$ را محاسبه کن.
- ۱۸: H_r را طبق (۱۰.۵) محاسبه کن.
- ۱۹: پایان شرط
- ۲۰: پایان حلقه درونی
- ۲۱: قرار ده $X_{k+1} = x_i$ که $i \in \{0, \dots, m-1\}$ به صورت تصادفی انتخاب می شود.
- ۲۲: پایان حلقه بیرونی

لم ۱.۳.۵. [۹۸] فرض می کنیم f تابعی λ -محدب قوی و Λ -هموار است و هر تابع f_i برای $i \in \{1, \dots, n\}$ تابعی محدب و دو بار مشتق پذیر پیوسته است. در این صورت، ثابت های $\gamma < \Gamma$ وجود دارند به طوری که

$$\gamma I \preceq H_r \preceq \Gamma I$$

برای هر $r \geq 1$. در [۹۸] ثابت شده است که γ و Γ دارای مقادیر

$$\gamma = \frac{1}{(d+M)\Lambda}, \quad \Gamma = \frac{((d+M)\Lambda)^{d+M-1}}{\lambda^{d+M}}$$

هستند.

لم ۲.۳.۵. [۹۸] فرض کنید X^* کمینه‌کننده یکتای تابع f باشد و $\mu_k = \nabla f(X_k)$. همچنین، فرض کنید $v_t = \nabla f_S(x_t) - \nabla f_S(X_k) + \mu_k$ گرادیان تصادفی با واریانس کاهش یافته باشد. آن‌گاه داریم

$$\mathbb{E}[\|v_t\|^2] \leq 4\Lambda(f(x_t) - f(X^*) + f(X_k) - f(X^*))$$

قضیه ۱.۳.۵. (همگرایی). [۹۸] فرض می‌کنیم f تابعی λ - محدب قوی و Λ - هموار است و هر تابع f_i برای $i \in \{1, \dots, n\}$ تابعی محدب و دو بار مشتق‌پذیر پیوسته است. به‌علاوه X^* را به عنوان جواب بهینه یکتای تابع در نظر می‌گیریم. همچنین، فرض می‌کنیم $\eta < \gamma\lambda/(2\Gamma^2\Lambda^2)$ و m را به قدری بزرگ در نظر می‌گیریم که در رابطه

$$\gamma\lambda > \frac{1}{2m\eta} + 2\eta\Gamma^2\Lambda^2$$

صدق کند. در این صورت برای هر $k \geq 0$ داریم

$$\mathbb{E}[f(X_k) - f(X^*)] \leq \alpha^k \mathbb{E}[f(X_0) - f(X^*)]$$

که پارامتر α به صورت

$$\alpha = \frac{1/(2m\eta) + \eta\Gamma^2\Lambda^2}{\gamma\lambda - \eta\Gamma^2\Lambda^2} < 1$$

در نظر گرفته می‌شود.

۴.۵ نتایج عددی

در این بخش، نتایج عددی حاصل از حل مساله مکان‌یابی توافقی با نرم ℓ_p با الگوریتم ارائه شده در بخش قبل را روی مثال‌های با مقیاس متفاوت ارائه می‌دهیم. همچنین نتایج حاصل از این روش را با روش‌های SVRG و SGD مقایسه می‌کنیم و برای درک بهتر نتایج، نمودارهای مرتبط را ارائه می‌دهیم. تمامی الگوریتم‌ها با استفاده از نرم افزار Matlab- R2017b و با یک رایانه همراه با پردازنده Intel core i7-4510U و مقدار حافظه ۸ گیگابایت پیاده‌سازی شده‌اند. برای هر سه الگوریتم پارامتر λ برابر 10^{-1} در نظر گرفته شده‌اند. برای روش‌های SVRG و SLBFGS طول گام ثابت 10^{-4} و برای روش SGD، طول گام نزولی $1/\|\nabla f(X_k)\|$ استفاده شده است.^{۴۲}

توجه می‌کنیم گرچه مساله مکان‌یابی توافقی با نرم ℓ_p محدب قوی نیست اما نتایج عددی ارائه شده در این بخش نشان می‌دهد که الگوریتم SLBFGS حتی برای این نوع مسایل نیز دارای عملکرد بسیار مطلوبی است.

شکل ۱.۵ به مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با نرم ℓ_2 برای مثال ۳۰ عضوی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها می‌پردازد. همان‌طور که در شکل می‌بینیم الگوریتم SGD مقادیر تابع هدف را به صورت نوسانی به دست آورده است و عملکرد چندانی مناسبی نداشته است اما الگوریتم SLBFGS و SVRG عملکرد مطلوبی داشته‌اند و مقادیر تابع هدف را به خوبی کاهش داده‌اند. به علاوه SLBFGS در تعداد تکرار کمتری به مقدار تابع هدف کمتری نسبت به SVRG دست یافته است.

در شکل ۲.۵ مقادیر واریانس برآوردگر گرادیان برای روش‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۳۰ عضو نمایش داده شده است. مطابق این شکل واریانس برآوردگر گرادیان در هر دو روش با سرعت بالایی کاهش می‌یابد که بیانگر عملکرد مطلوب هر دو روش در کاهش واریانس است.

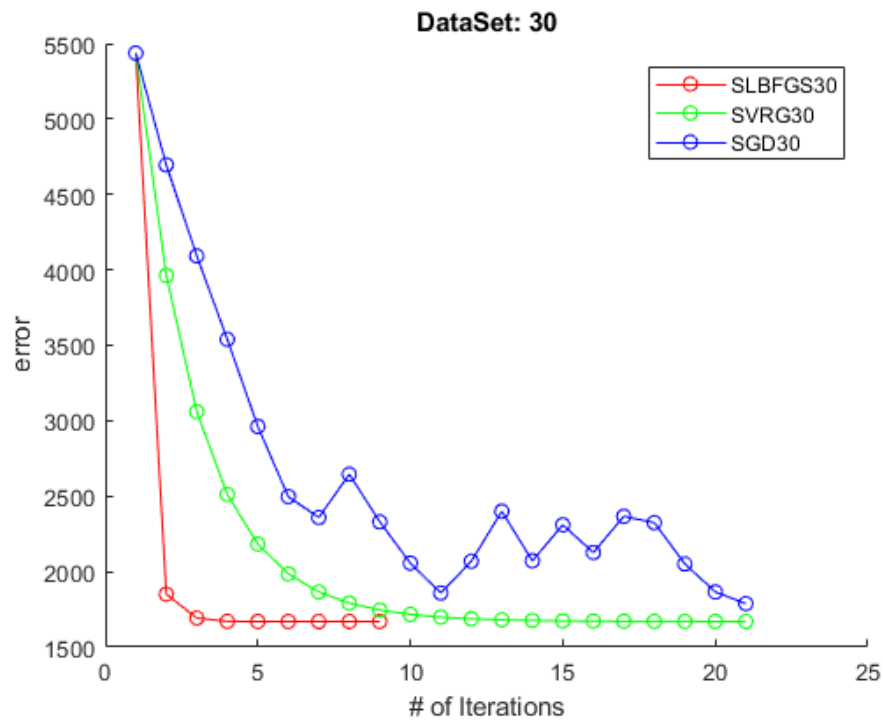
همان‌طور که در شکل ۳.۵ مشاهده می‌کنیم، مقادیر تابع هدف برای مجموعه مثال دارای ۱۰۰ عضو در روش‌های SVRG و SLBFGS با سرعت بالایی کاهش یافته‌اند در حالی که روش SGD به مقادیر تابع هدف خوبی دست نیافته است و بسیار کند پیش می‌رود.

مشابه مجموعه مثال ۳۰ عضوی، برای مجموعه مثال ۱۰۰ عضوی نیز واریانس برآوردگر گرادیان با سرعت بالایی کاهش یافته است که این مطلب در شکل ۴.۵ قابل مشاهده است.

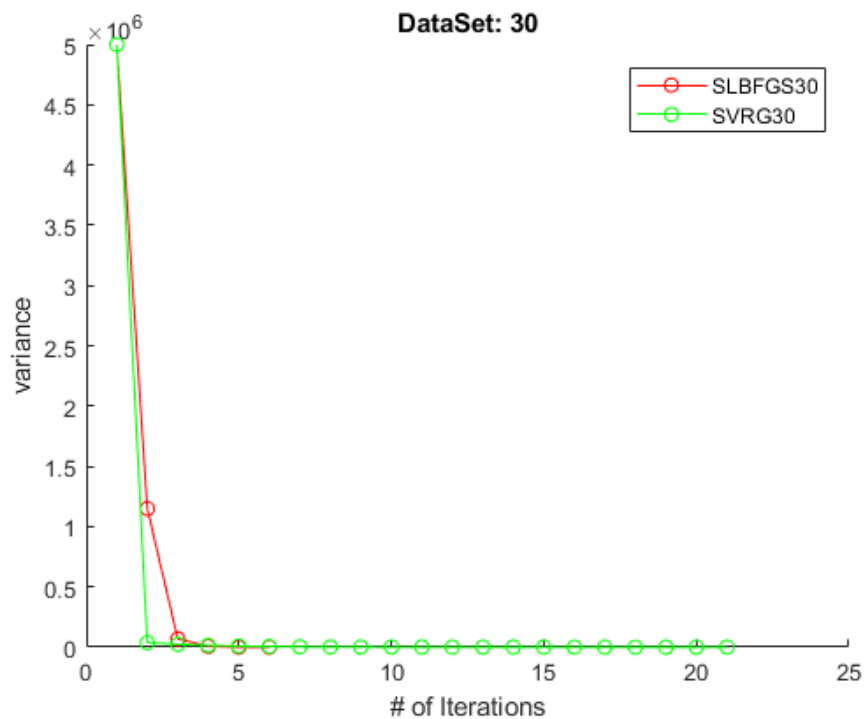
شکل ۵.۵ به مقایسه مقادیر تابع هدف بر حسب تکرارها با استفاده از سه روش ذکر شده می‌پردازد. در این شکل نیز روش SLBFGS عملکرد بسیار مطلوبی از خود نشان می‌دهد که حاکی از کاهش واریانس برآوردگر گرادیان می‌باشد.

در شکل ۶.۵ نیز کاهش واریانس به خوبی قابل مشاهده است با این تفاوت که در تکرار

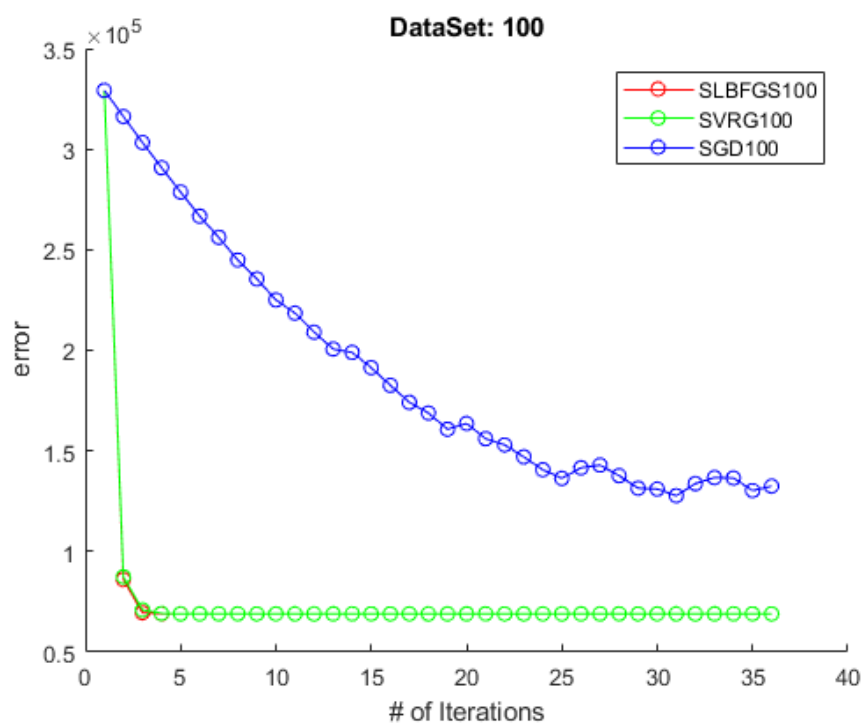
^{۴۲} به منظور انتخاب طول گام مناسب، چندین طول گام متفاوت را آزمون کرده و بهترین طول گام ممکن را انتخاب کرده‌ایم.



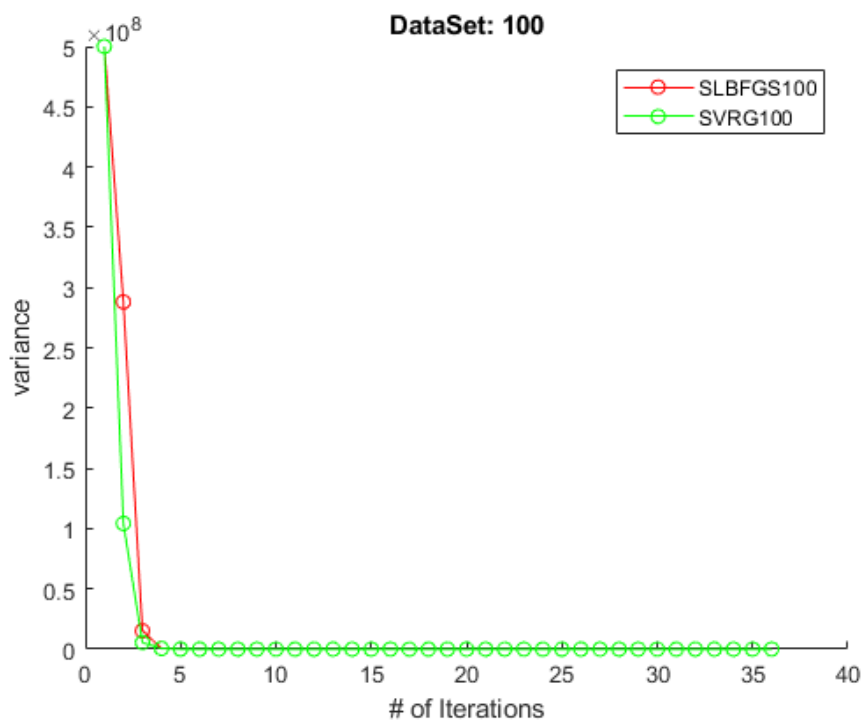
شکل ۱.۵: مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۳۰ عضو



شکل ۲.۵: مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۳۰ عضو

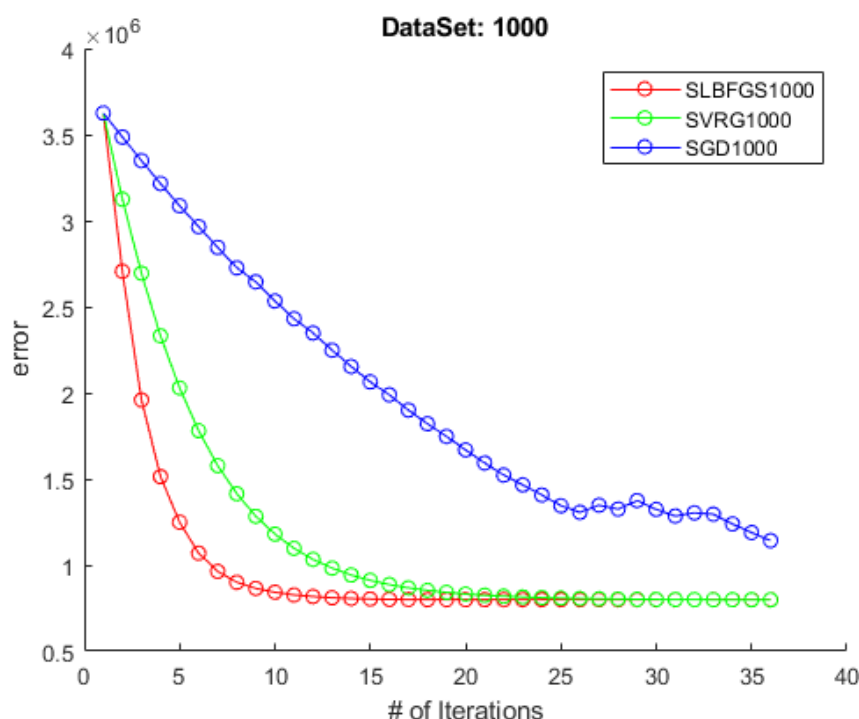


شکل ۳.۵: مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰ عضو



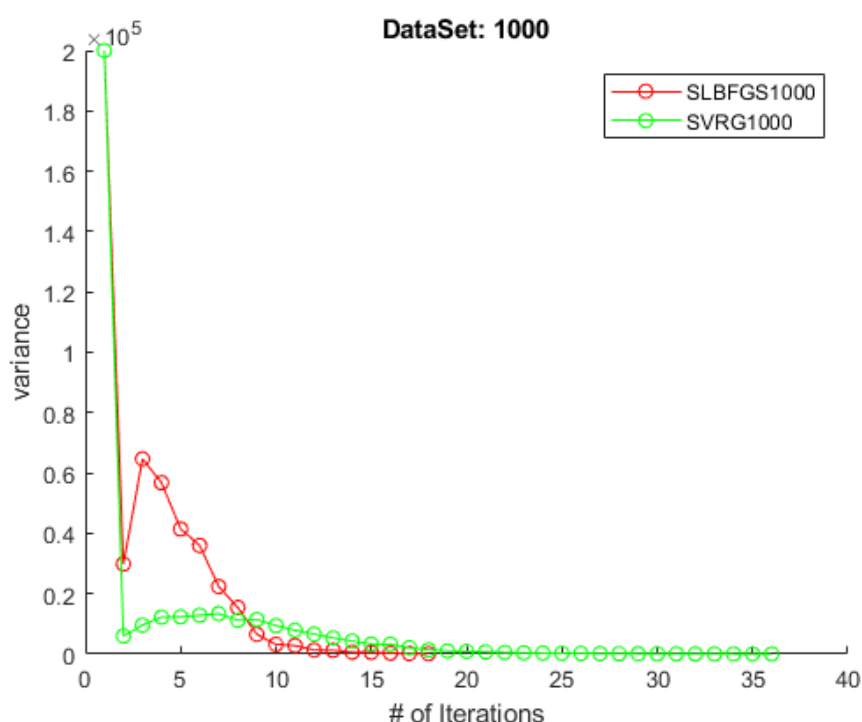
شکل ۴.۵: مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰ عضو

چهارم به دلیل تصادفی بودن روش و نمونه‌گیری نامطلوب، واریانس برآوردگر گرادیان افزایش یافته است اما مجدداً به روند کاهشی خود ادامه داده است و واریانس حاصل از روش SLBFGS در تعداد تکرار کمتری به میزان کمتری دست یافته است که موجب سرعت بالاتر روش در کاهش مقدار تابع هدف شده است که این مطلب نتایج حاصل در شکل ۵.۵ را نیز تایید می‌کند.



شکل ۵.۵: مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو

در جدول ۲.۵ مقادیر بهینه، جواب بهینه و سرعت اجرای الگوریتم‌های SGD، SVRG و SLBFGS برای مقادیر مختلف p ارائه شده است.



شکل ۶.۵: مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو

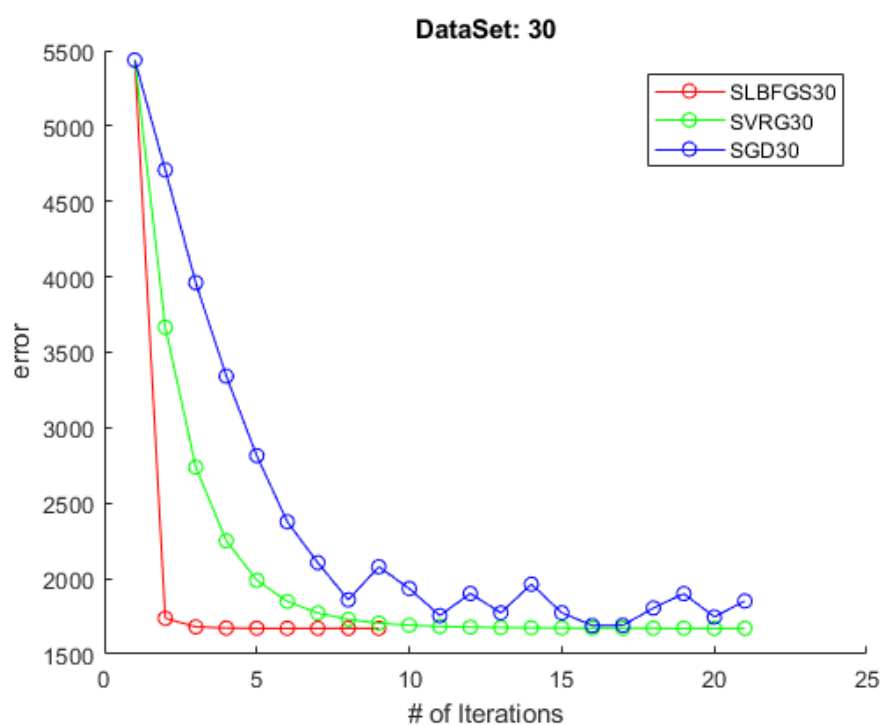
جدول ۲.۵: مقادیر بهینه، جواب بهینه و سرعت اجرای الگوریتم‌های SGD، SVRG و SLBFGS برای مقادیر مختلف p

مقادیر p	اندازه مجموعه مثال	الگوریتم	مقدار بهینه	جواب بهینه	زمان اجرا
$p=2$	$n=30$	SGD	۱۸۵۴/۹	(۷/۲۷۷۲, ۶/۱۰۴۰)	۰/۰۳۰۷
		SVRG	۱۶۶۹/۵	(۸/۲۳۳۶, ۷/۶۵۰۱)	۰/۰۷۰۰
		SLBFGS	۱۶۶۹/۲	(۸/۲۷۷۱, ۷/۶۹۰۲)	۰/۰۷۱۵
$p=2$	$n=100$	SGD	$1/3188 \times 10^5$	(۱۵/۷۴۳۸, ۱۶/۱۶۸۶)	۰/۰۴۶۷
		SVRG	$0/6900 \times 10^5$	(۳۰/۴۵۸۴, ۲۹/۲۰۰۵)	۰/۳۷۹۳
		SLBFGS	$0/6900 \times 10^5$	(۳۰/۴۶۴۲, ۲۹/۲۰۷۷)	۰/۲۸۷۷
$p=2$	$n=1000$	SGD	$1/1183 \times 10^6$	(۲۳/۲۴۸۰, ۱۹/۱۳۴۷)	۰/۲۵۷۶
		SVRG	$0/7983 \times 10^6$	(۳۰/۲۳۵۷, ۳۰/۳۷۱۳)	۲/۴۶۵۷
		SLBFGS	$0/7992 \times 10^6$	(۲۹/۹۶۹۲, ۳۰/۰۷۴۶)	۱/۴۲۶۵
$p=1.5$	$n=30$	SGD	۲۹۳۵/۰	(۶/۹۹۷۰, ۳/۸۸۸۸)	۰/۱۵۹۱
		SVRG	۲۰۳۵/۳	(۸/۲۳۳۶, ۷/۶۵۰۱)	۰/۱۱۴۹
		SLBFGS	۲۰۳۵/۹	(۸/۲۷۷۲, ۷/۶۹۰۲)	۱/۱۸۶۸
$p=1.5$	$n=100$	SGD	$1/5111 \times 10^5$	(۱۹/۱۸۴۲, ۱۴/۰۴۱۲۸)	۰/۱۲۴۴
		SVRG	$0/8457 \times 10^5$	(۲۹/۸۹۲۹, ۲۹/۳۸۸۴)	۱/۳۷۶۶
		SLBFGS	$0/8458 \times 10^5$	(۳۰/۰۵۳۳, ۲۹/۵۲۲۱)	۱/۶۱۴۵
$p=1.5$	$n=1000$	SGD	$1/3037 \times 10^6$	(۲۰/۶۶۵۵, ۲۲/۴۷۸۲)	۱/۰۱۱۸
		SVRG	$0/9837 \times 10^6$	(۳۰/۱۹۷۰, ۳۰/۵۵۷۸۱)	۶/۴۳۷۱
		SLBFGS	$0/9835 \times 10^6$	(۳۰/۳۲۳۹, ۳۰/۷۷۶۲)	۳/۴۸۳۲

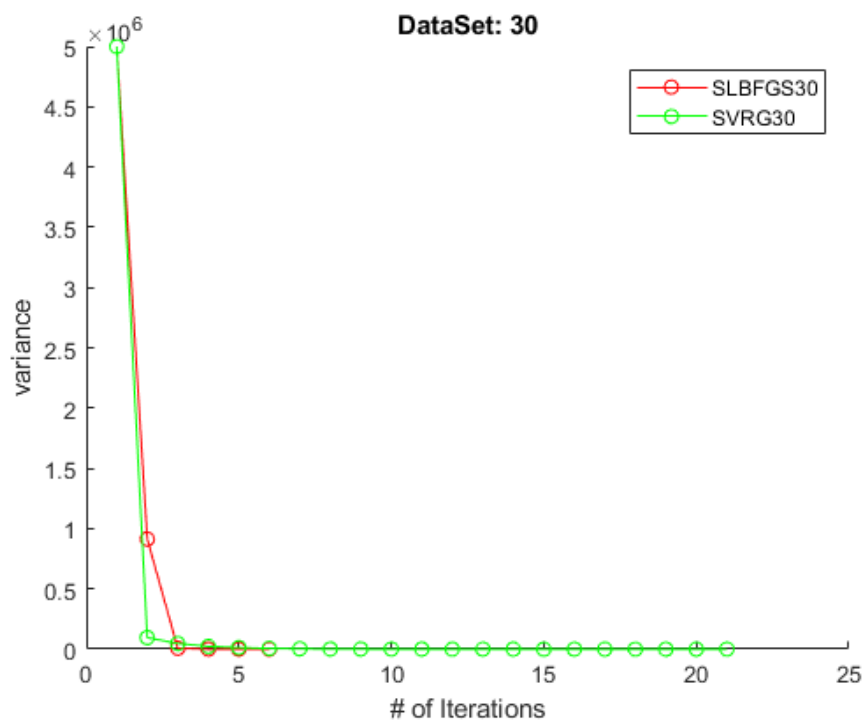
در روش‌های بهینه‌سازی اندازه نمونه تصادفی نقش به‌سزایی در کاهش واریانس برآوردگر و در نتیجه عملکرد الگوریتم ایفا می‌کند. بنابراین، الگوریتم‌های SGD، SVRG و SLBFGS را با اندازه نمونه تصادفی مختلف اجرا کرده‌ایم. اندازه نمونه تصادفی که همان تعداد هسی‌ها و گرادیان‌های تصادفی برای محاسبه جهت جستجو است را با $m = |S|$ نمایش می‌دهیم. شکل‌های ۷.۵ تا ۱۲.۵ نمودارهای الگوریتم‌های مذکور برای مقادیر مختلف m را نمایش می‌دهند که بیانگر این مطلب هستند که افزایش اندازه نمونه تصادفی باعث کاهش واریانس برآوردگر گرادیان شده و عملکرد الگوریتم‌ها را بهبود می‌بخشد. در جدول ۳.۵ مقادیر بهینه، جواب بهینه و سرعت اجرا و واریانس برآوردگر گرادیان الگوریتم‌های SGD، SVRG و SLBFGS برای مقادیر مختلف m ارائه شده است.

جدول ۳.۵: مقادیر بهینه، جواب بهینه، سرعت اجرا و واریانس برآوردگر گرادیان الگوریتم‌های SGD، SVRG و SLBFGS برای مقادیر مختلف m

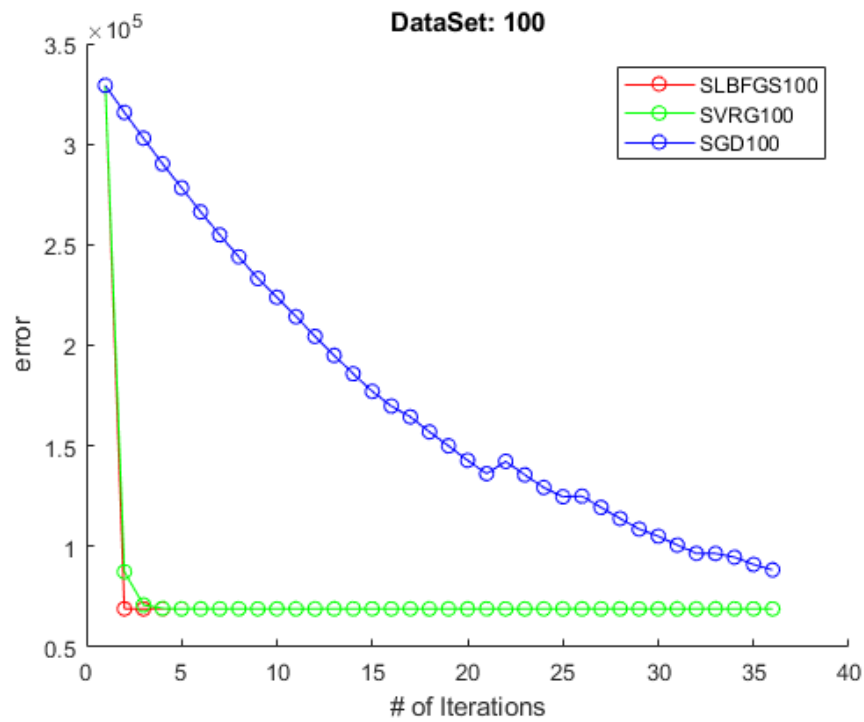
مقادیر m	اندازه مجموعه مثال	الگوریتم	مقدار بهینه	جواب بهینه	زمان اجرا	واریانس
$m=4$	$n=30$	SGD	$3/1523 \times 10^3$	($4/4316, 4/0064$)	$0/0095$	—
		SVRG	$1/6709 \times 10^3$	($8/2300, 7/5365$)	$0/1720$	$0/5765$
		SLBFGS	$1/6703 \times 10^3$	($8/2612, 7/5637$)	$0/4108$	$5/8282 \times 10^3$
$m=4$	$n=100$	SGD	$1/2717 \times 10^5$	($18/2118, 14/7525$)	$0/0522$	—
		SVRG	$0/6900 \times 10^5$	($30/4584, 29/2005$)	$0/3072$	$5/6211 \times 10^{-22}$
		SLBFGS	$0/6900 \times 10^5$	($30/4642, 29/2077$)	$0/0661$	$2/8786$
$m=4$	$n=1000$	SGD	$1/3202 \times 10^6$	($20/2169, 16/5848$)	$0/2543$	—
		SVRG	$0/7983 \times 10^6$	($30/2357, 30/3713$)	$1/5258$	$24/1414$
		SLBFGS	$0/7984 \times 10^6$	($30/1902, 30/3190$)	$1/4853$	$100/4597$
$m=10$	$n=30$	SGD	$1/8520 \times 10^3$	($6/8308, 6/5223$)	$0/0066$	—
		SVRG	$1/6709 \times 10^3$	($8/2298, 7/5364$)	$0/1637$	$0/5664$
		SLBFGS	$1/6707 \times 10^3$	($8/2341, 7/5429$)	$0/3625$	$6/4919$
$m=10$	$n=100$	SGD	$0/8835 \times 10^5$	($24/2827, 20/1608$)	$0/0599$	—
		SVRG	$0/6900 \times 10^5$	($30/4584, 29/2005$)	$0/2345$	$5/0009 \times 10^{-22}$
		SLBFGS	$0/6900 \times 10^5$	($30/4584, 29/2005$)	$0/0473$	$0/003145$
$m=32$	$n=1000$	SGD	$0/9340 \times 10^6$	($23/9485, 24/6625$)	$0/2624$	—
		SVRG	$0/7983 \times 10^6$	($30/2325, 30/3679$)	$1/4199$	$22/9194$
		SLBFGS	$0/7901 \times 10^6$	($30/2402, 30/3200$)	$15/4827$	$94/7234$



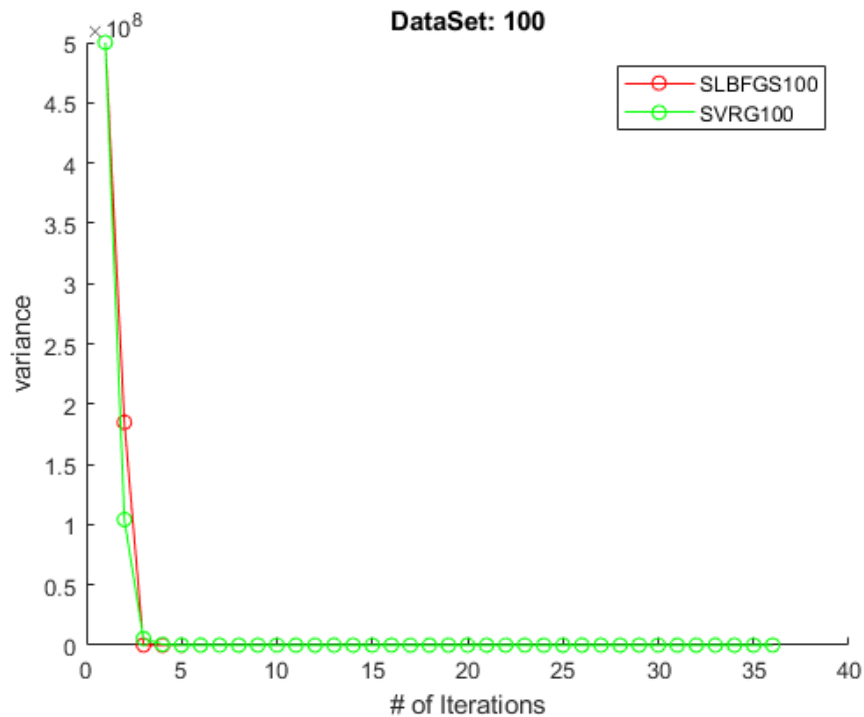
شکل ۷.۵: مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با $m = 10$ و 30 عضو



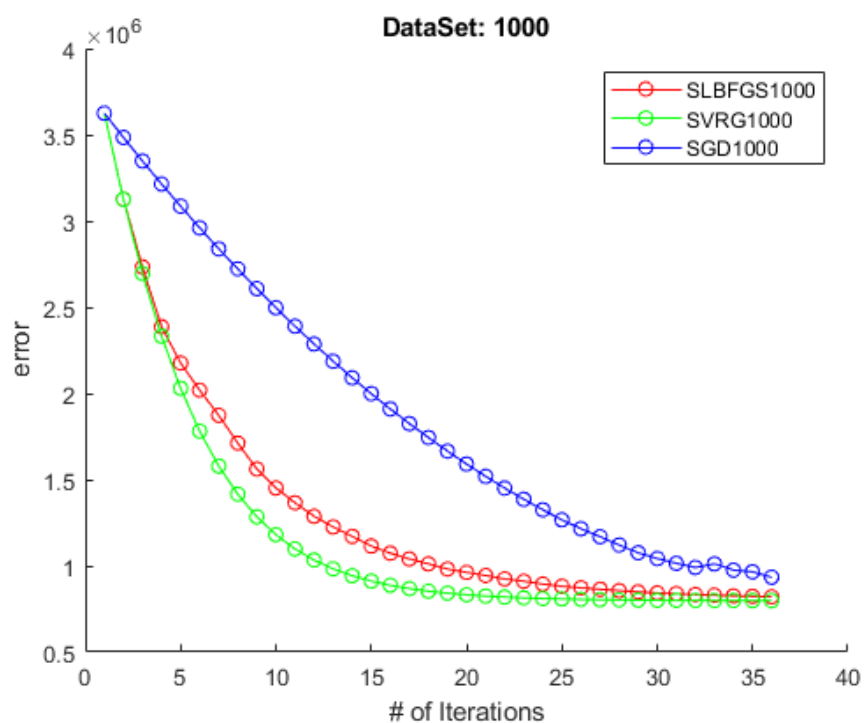
شکل ۸.۵: مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با $m = 10$ و 30 عضو



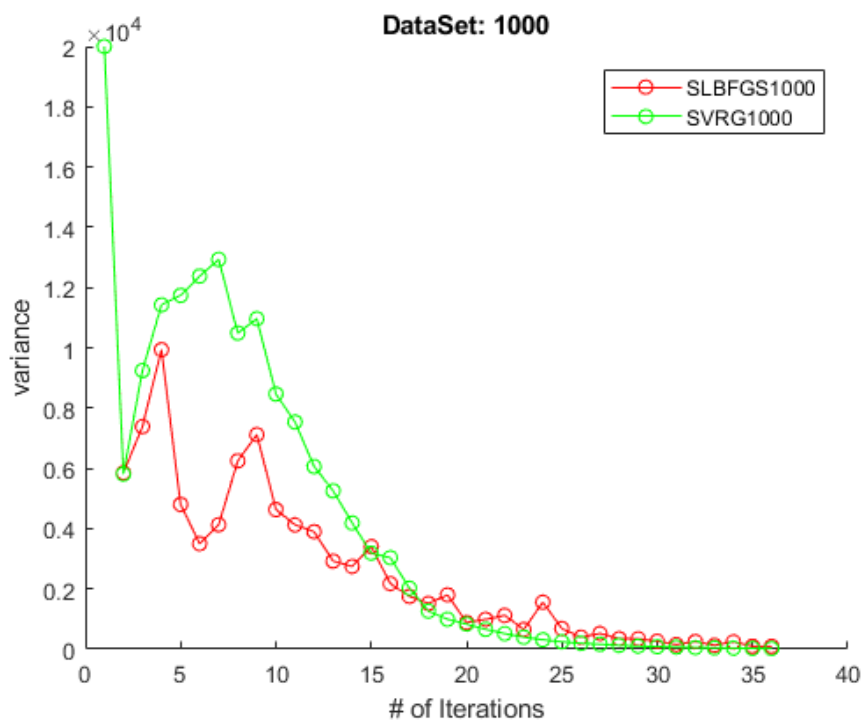
شکل ۹.۵: مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰ عضو و $m = 10$



شکل ۱۰.۵: مقایسه مقادیر واریانس برآوردگر گرادینت مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰ عضو و $m = 10$



شکل ۱۱.۵: مقایسه مقادیر تابع هدف مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SGD، SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو و $m = 32$



شکل ۱۲.۵: مقایسه مقادیر واریانس برآوردگر گرادیان مساله مکان‌یابی توافقی با استفاده از الگوریتم‌های SVRG و SLBFGS برحسب تکرارها برای مثال با ۱۰۰۰ عضو و $m = 32$

نتیجه‌گیری و پیشنهادات آینده

در این رساله به بررسی مسائل مکان‌یابی توافقی تک‌وسیله‌ای تحت توابع زیان متقارن و نامتقارن پرداختیم و به کمک روش‌های بهینه‌سازی غیرخطی توابع زیان مربوطه را کمینه کرده و نتایج حاصل را ارائه دادیم. هم‌چنین، به کمک روش‌های بهینه‌سازی تصادفی به حل مسایل مکان‌یابی توافقی مقیاس بزرگ پرداختیم و نتایج عددی حاصله را ارائه دادیم. نتایج حاصل از پیاده‌سازی روش‌های مختلف بهینه‌سازی غیرخطی نشان می‌دهند که روش‌های اصلاح‌شده BFGS نسبت به دیگر روش‌های مطرح‌شده دارای خصوصیات همگرایی مطلوبی هستند. هم‌چنین روش‌های بهینه‌سازی تصادفی برای مسائل مقیاس بزرگ عملکرد خوبی به دست آوردند و در زمان اجرای کمتری به مقدار تابع هدف کمتری دست یافتند.

به عنوان اهداف و پیشنهادات آینده می‌توان به حل مسائل مکان‌یابی چندوسیله‌ای توافقی با استفاده از روش‌های بهینه‌سازی غیرخطی پرداخت. هم‌چنین، برای مسایل مقیاس بزرگ می‌توان رویکردی تصادفی در نظر گرفت و از تکنیک‌های متنوع تصادفی بهره برد. به‌علاوه می‌توان توابع زیان متفاوتی که دارای خصوصیات دیگری هستند را در نظر گرفت و به کمینه‌سازی آن‌ها پرداخت. به‌علاوه، حوزه بهینه‌سازی تصادفی تکنیک‌های بسیار گسترده‌ای دارد که یکی از آن‌ها تکنیک اندازه نمونه پویا نام دارد. در این تکنیک اندازه نمونه تصادفی با استفاده از مقدار واریانس برآوردگر در هر تکرار به صورت پویا تغییر می‌کند تا بتوان به کاهش بیشتر واریانس کمک نمود و سرعت اجرای الگوریتم را افزایش داد. هدف دیگری که در مطالعات آتی به آن خواهیم پرداخت، حل مسایل مکان‌یابی توافقی مقید است که در این مسایل قیودی به مساله در نظر گرفته می‌شود.

مراجع

- [۱] مرادی ح.، شتاب بوشهري ن.، کورنک بهشتي ع.، پورزاهدي ح.، مکان یابی مراکز ارائه خدمات رقابتي با هدف کاهش ازدحام ترافیک شهري مطالعه موردی: مکان یابی مراکز سلامت در شهر اصفهان، مجله علمي – پژوهشي مدیریت تولید و عملیات، سال اول، شماره اول، پاییز و زمستان ۱۳۸۹، ص. ۳۱-۵۲.
- [۲] سهامی ح.، آمایش و مکان یابی، (۱۳۸۶) چاپ اول، دانشگاه صنعتی مالک اشتر، پژوهشکده مهندسی پدافند غیر عامل.
- [3] Abbasi, F., Tabriz, A. A. (2017), "Selection of bank branches location based on rough set theory –multi choice goal programming", **Modern Researches in Decision Making**, 2 (1), pp 119-148.
- [4] Agarwal, N. , Allen-Zhu, Z. , Bullins, B. , Hazan, E. , and Tengyu Ma.(2017), "Finding Approximate Local Minima for Nonconvex Optimization in Linear Time", **In STOC**.
- [5] Agarwal, N. , Bullins, B. , and Hazan, E. , (2017), "Second-order stochastic optimization for machine learning in linear time", **J. Mach. Learn. Res.**, 18, pp. 4148–4187.
- [6] Allen-Zhu, Z. (2018), "Katyusha: The first direct acceleration of stochastic gradient methods", **Journal of Machine Learning Research**, 18, pp 1-51.
- [7] Amiri, M., Taghavi Fard, M. T., Aghaei M. (2016), " Development of three-objective model for the location–allocation of assistance centers in a probabilistic condition of availability to emergency vehicles", **Modern Researches in Decision Making**, 1 (2), pp 1-27.
- [8] Andrei N. (2018). "A double parameter scaled BFGS method for unconstrained optimization", **Journal of Computational and Applied Mathematics**, 332, pp 26–44.

- [9] Andrei N. (2018). "An adaptive scaled BFGS method for unconstrained optimization", **Numerical Algorithms**, 77, pp 413–432.
- [10] Arashi, M. , Tabatabaey, S. M .M. and Khan, S. , (2008), "Estimation in multiple regression model with elliptically contoured errors under MLINEX loss", **Journal of Applied Probability and Statistics**, 3, pp 23-35.
- [11] Babaie-Kafaki S. (2016). "A modified scaling parameter for the memoryless BFGS updating formula", **Numerical Algorithms**, 72, pp 425–433.
- [12] Babaie-Kafaki S., Aminifard Z. (2019). "Two-parameter scaled memoryless BFGS methods with a nonmonotone choice for the initial step length", **Numerical Algorithms**, 82, pp 1345–1357.
- [13] Babaie-Kafaki, S.(2013), "A new proof for the sufficient descent condition of Andrei's scaled conjugate gradient algorithms," **Pac. J. Optim.** 9(1), pp 23–28.
- [14] Babaie-Kafaki, S.(2015), "On optimality of the parameters of self-scaling memoryless quasi-Newton updating formulae," **J. Optim. Theory Appl.** 161(1), pp 91-101.
- [15] Babaie-Kafaki, S.,(2012) "A note on the global convergence theorem of the scaled conjugate gradient algorithms proposed by Andrei," **Comput. Optim. Appl.** 52(2), pp 409–414.
- [16] Balcik, B., and Beamon, B. M., (2008), "Facility location in humanitarian relief", **International Journal of Logistics Research and Applications**, 11 (2), pp 101 - 121 .
- [17] Barbati M., Piccolo C.(2015), "Equality measures properties for location problems", **Optimization Letters**, 10, pp 903-920.
- [18] Bashiri, M., Garmeyi, Y. (2014), "Solving Multi Criteria Gradual Covering Problem Using Simulated Annealing and Artificial Neural Network". **IQBQ**, 17 (4), pp 25-41.
- [19] Bazaraa, M.S., Sherali H., Jarvis J., (2005), "**Linear programming and network flows**", third ed. John Wiley and Sons.
- [20] Beresnev, V., (2013), "Branch-and-bound algorithm for a competitive facility location problem", **Computers and Operations Research**, 40 (8), pp 2062-2070.

- [21] Berger, J. O. , (1980),” **Statistical Decision Theory**”, Springer, NewYork.
- [22] Blanquero R., and Carrizosa E., (2009), ”Continuous location problems and Big Triangle Small Triangle: constructing better bounds”, **Journal of Global Optimization**, 45 (3), pp 389-402.
- [23] Bottou, L. , Curtis, F. E. and Nocedal, J. , (2018), ”Optimization methods for large-scale machine learning”, **SIAM Rev.**, vol. 60, no. 2, pp. 223-311.
- [24] Brimberg, J. (1995), ”The Fermat-Weber location problem revisited”, **Mathematical Programming**, 71, pp 71-76.
- [25] Brimberg, J., Love, R., Mladenović, N. (2009), ”Extension of the Weiszfeld procedure to a single facility minisum location model with mixed lp norms”, **Math. Meth. Oper. Res.**, 70, pp 269-283
- [26] Brotcorne L., Laporte G., and Semet F., (2002), ”Fast heuristics for large scale covering location problems”, **Computers and Operations Research**, 29, 651 - 665.
- [27] Byrd, R. H. , Gillian M. Chin, Will Neveitt, and Jorge Nocedal(2011), ”On the use of stochastic Hessian information in optimization methods for machine learning”, **SIAM Journal on Optimization**, 21(3), pp 977-995.
- [28] Byrd, R. H. , Hansen, S. L. , Nocedal, J. , and Singer, Y.(2016), ”A stochastic quasi-Newton method for large-scale optimization”, **SIAM Journal on Optimization**, 26 (2), pp 1008-1031.
- [29] Cabello S., Díaz-Báñez J. M., Langerman S., Seara C., and Ventura I., (2010), ”Facility location problems in the plane based on reverse nearest neighbor queries”, **European Journal of Operational Research**, 202, pp 99–106 .
- [30] Campbell J. F., (1994), ”Integer programming formulations of discrete hub location problems”, **European Journal of Operational Research**, 72,pp 387 - 405.
- [31] Cappanera P., (1999), ”A survey on obnoxious facility location problems”, Technical Report: TR-99-11, **Department of Computer Science**, University of Piza.
- [32] Charch R.L., Garfinkel R.S., (1978), ”Location an obnoxious facility on a network”, **Trans. Sci.**, 12, pp 107-118.

- [33] Chen, R. (2011), "Noniterative solution of some Fermat-Weber location problems", **Advances in Operations Research**, 10 pages.
- [34] Cheriyan J., Ravi R., (1998), "**Aproximation for Networks Problems**".
- [35] Contreras I., Fernández E., (2012), "General network design: A unified view of combined location and network design problems", **European Journal of Operational Research**, 219, pp 680-697.
- [36] Corra E.S., Steiner M.T.A. Freitas A.A., Cornieri C., (2001), "**A genetic algorithm for the p-median problem**", Available in Address: Cite-seer.nj.ncg.com/458099.html.
- [37] Current J., Daskin M., Schilling D., (1929), "Discrete network location models", Forthcoming as Chapter 3 in: Facility Location Theory and Methods. Z.Drezner and H.Hamacher (eds.), 2001. **Chicago University Press, Chicago, Illinois**.
- [38] Daskin, M. S., (1995), "**Network and Discrete Location: Models, Algorithms and Applications**", Wiley, NewYork.
- [39] Defazio, A. , Bach, F. , and Simon Lacoste-Julien. (2014) "Saga: A fast incremental gradient method with support for non-strongly convex composite objectives", **In Advances in Neural Information Processing Systems**, pp 1646-1654.
- [40] Dibble C., Densham P.J., (1993), "**Generating interesting alternatives in GIS and SDSS using genetic algorithms**", GIS/LIS Symposium, University of Nebraska, Lincoln.
- [41] Dolan, E.D., Moré, J.J. (2002): "Benchmarking optimization software with performance profiles", **Math. Program.**, 91, pp 201–213.
- [42] Dominik Jena S., Cordeau J. F., and Gendron B. , (2017), "Lagrangian Heuristics for Large-Scale Dynamic Facility Location with Generalized Modular Capacities", **INFORMS Journal on Computing**, 29 (2), pp 388 - 404 .
- [43] Drezner Z., and Wesolowsky G. O., (1978), "Facility location on a sphere", **The Journal of the Operational Research Society**, 29 (10), pp 997-1004.
- [44] Ebery J., (2001), "Solving large single allocation p-hub problems with two or three hubs", **European Journal of Operational Research**, 128, pp 447-458.

- [45] Erdogdu, M. A. and Montanari, A. .(2015), "Convergence rates of sub-sampled Newton methods", **In Advances in Neural Information Processing Systems**, pp 3034-3042.
- [46] Erkut E., Bozkaya B., Zhang J., (2001), "**An effective genetic algorithm for the p-median problem**".
- [47] Ernst A. T. , and Hrishnamoorthy M., (1998), "Exact and heuristic algorithms for the uncapacitated multiple allocation p- hub median problem", **European Journal of Operational Research**, 104, pp 100-112.
- [48] Eyster, J.W., White, J.A., Wierwille, W.W. (1973), "On solving multifacility location problems using a hyperboloid approximation procedure", **AIIE Transacions**, 5,pp 1–6.
- [49] Fathali J., (2006), "A genetic algorithm for the p-median problem with pos/neg weights", **Applied Mathematics and Computation**, 183, pp 1071-1083.
- [50] Fathali J., and Taghizadeh Kakhki H., (2006), "Solving the p-median problem with pos/neg weights by variable neighborhood search and some results for special cases", **European Journal of Operational Research**, 170, pp 440-462.
- [51] Fathali, J. (2014), "Backup multifacility location problem with norm", **OPSEARCH**, 52, pp 382-391.
- [52] Fathali, J., Jamalian, A. (2017), "Efficient methods for goal square Weber location problem", **Iranian Journal of Numerical Analysis and Optimization**, 7, pp 65-82.
- [53] Fathali, J., Zaferanieh, M. and Nezakati, A. (2009), "A BSSS algorithm for the location problem with minimum square error", **Advances in Operations Research**.
- [54] Francis, R. L. , (1964), "On the location of multiple new facilities with respect to existing facilities", **The Journal of Industrial Engineering**, 15, pp 10 - 107 .
- [55] Francis R. L., McGinnis Jr.L.F., White J.A., (1992), "**Facility Layout and Location: An Analytical Approach**", Prentice Hall.
- [56] Galinier P., and Hertz A., (2007), " Solution techniques for the large set covering problem", **Discrete Applied Mathematics**, 155, pp 312-326.

- [57] Garey M.R., Johnson D.S., (1979), "**Computers and intractability: A guid to the theory of NP-Completeness**", Freeman, San Francisco, California.
- [58] Gargari, A., Lucas, E. C., (2007)"Imperialist competitive algorithm: an algorithm for optimization inspired by imperialist competitive", **IEEE Congress on Evolutionary com.**
- [59] Goldman, A.J., Dearing, P.M., (1975), "Concepts of optimal location for partially noxious facilities", **ORSA Bull.** 23, 1, B31.
- [60] Gorbunov, E., Hanzely, F., and Richtarik, P. A. (2020) "Unified Theory of SGD: Variance Reduction, Sampling, Quantization and Coordinate Descent", **Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics**, PMLR 108, pp 680-690.
- [61] Gower, R. M., Loizou, N., Qian, X., Sailanbayev, A., Shulgin, E., and Richtárik, P. (2019), "SGD: General analysis and improved rates", **arXiv preprint arXiv:1901.09401**.
- [62] Haddadi S., (1997), " Simple lagrangian heuristic for the set covering problem", **European Journal of Operational research**, 99, pp 200-204.
- [63] Hakimi, S.L., (1964), "Optimum location of switching centers and the absolute centers and medians of a graph", **Oper. Res**, 12, pp 450-459.
- [64] Hamacher H. W., Nickel S., (1998), "classification of location models", **Location Sci.**, 6, pp 229-242.
- [65] Handler G. Y., Mirchandani P. B., (1979), **Location on Networks: Theory and Algorithms**, MIT Press, Cambridge.
- [66] Handler G. Y. , (1973), Minimax location of a facility in a undirected tree network, **Transportation Science**, 7, pp 287-293.
- [67] Hansen P., and Mladenovic N., (1997), " Variable neighborhood search for the p-median", **Location Science**, 5 (4), pp 207-226.
- [68] Hansen P., Labbe M., Peeters D., Thisse J.F., (1987), "Single facility location on networks", **Annals of Discrete Math**, 3, pp 113-146.

- [69] Hedetniemi S. M. , Cockayne E. J., and Hedetniemi S. T., (1981), "linear algorithm for finding the Jordan center and path center of a tree", **Trancs.SCI**, 15, pp 98-114.
- [70] Ho S. C., (2015), "An iterated tabu search heuristic for the Single Source Capacitated Facility Location Problem", **Applied Soft Computing**, 27, pp 169 - 178 .
- [71] Ho, Y. C., and Pepyne, D. L. (2001), "Simple explanation of the No free lunch theorem of optimization", **Proceeding of the 40th IEEE, Conf. on Decision and Control**, USA.
- [72] Honggzhong, J., Fernando, O., Maged, D., (2007) "A modeling framework for facility location of medical services for large-scale emergencies",**IIE Transactions**, 39: 1, pp 41-55.
- [73] Hosage, M., Goodchild M.F., (1986), "Discrete space location-allocation solutions from genetic algorithms", **Ann. Oper. Res**, 6, pp 35-46.
- [74] Iyigun, C., Ben-Israel, A (2010)., "A generalized Weiszfeld method for the multifacility location problem", **Oper. Res. Lett.**, 38, pp 207–214.
- [75] Jamalian, A. and Fathali, J. (2009), "Linear programming for the location problem with minimum absolute error",**World Applied Sciences Journal**, 7, pp 1423-1427.
- [76] Johnson, R. and Zhang, T. ,(2013), "Accelerating stochastic gradient descent using predictive variance reduction",**In Advances in Neural Information Processing Systems**, pp 315-323.
- [77] Kalcsics J., (2011), "The multi-facility median problem with pos/neg weights on general graphs", **Computers and Operations Research**, 38, pp 674 - 682 .
- [78] Karvin O., and Hakimi S.L., (1969), "An alghoritmie approach to network location problems; part 2, the p-medians", **SIAM Journal on Applied Mathematics**, 37, pp 539 - 560.
- [79] Karvin O., and Hakimi S. L., (1979), "An algorithm approach to network location problems. Part I: The p-centers", **SIAM Journal of Applied Mathematics**, 37, pp 513 - 538.
- [80] Klamroth K., (2001), "A reduction result for location problems with polyhedral barriers", **European Journal of Operational Research**, 130, pp 486 - 497 .

-
- [81] Ko M. T., and Lee R. C. T., (1991), "On weighted rectilinear 2-center and 3-center problems", **Information Sciences**, 54, pp 169-190.
 - [82] Kovalev, D. , Horváth, S. , Richtárik, P. (2020). "Don't Jump Through Hoops and Remove Those Loops: SVRG and Katyusha are Better Without the Outer Loop", Proceedings of the 31st International Conference on Algorithmic Learning Theory, PMLR 117, pp 451-467.
 - [83] Krarup J., Pruzan P. M., (1983), "The simple plant location problem: Survey and synthesis", **EJOR**, 12, pp 36-81.
 - [84] Kuhn, H.W. (1967), "On a pair of dual nonlinear programs", **Nonlinear programming, J., Abadie(ed.)**, North-Holland Publishing Co., Amesterdam.
 - [85] Li S., (2013), "A 1.488 approximation algorithm for the uncapacitated facility location problem", **Information and Computation**, 222, pp 45 - 58.
 - [86] Lin, H. , Mairal, J. , and Harchaoui, Z. (2015), "A universal catalyst for first-order optimization", **In Advances in Neural Information Processing Systems**, pp 3384-3392.
 - [87] Lin, Q. , Lu, Z. , and Lin Xiao. (2014), "An accelerated proximal coordinate gradient method", **In Advances in Neural Information Processing Systems**, pp 3059-3067.
 - [88] Love, R. F., Morris, J. D., and Wesolowsky, G. O. (1988), "Facilities locations: models and methods", **Publications in Operations Research Series**, vol. 7, North-Holland, Amsterdam.
 - [89] Maranzana F.E., (1964), "On the location of supply points to minimize transport costs", **Oper. Res Quarterly**, 15, pp 261-270.
 - [90] Martens, J. (2010), "Deep learning via Hessian-free optimization", In International Conference on Machine Learning, pp 735-742.
 - [91] Meyer, R. and Philip, M. (1972), "Modified Damped Least Squares: An Algorithm for Non-linear Estimation", **J. Inst. Maths Applies**, 9, pp 218-233.
 - [92] Miehle, W. (1958), "Link-length minimization in networks", **Oper. Res.**, 6, pp 232-243.

- [93] Minika E., (1958), "The optimal location of a path or tree in a tree network", **Networks**, 15, pp 309-321.
- [94] Mirchandani P. B, and Francis R. L., (1990), "**Discrete location theory**", Wiley, New York.
- [95] Mokhtari, A. and Ribeiro, A. (2014), "RES: regularized stochastic BFGS algorithm", **IEEE Transactions on Signal Processing**, 62(23), pp 6089-6104.
- [96] Montoya A., Velez-Gallego M. C., and Villegas J. G., (2016), "Multi-product capacitated facility location problem with general production and building costs", **NETNOMICS: Economic Research and Electronic Networking**, 17 (1), pp 47 - 70.
- [97] Moreno-Perez J.A., Moreno-Vega J.M., Meladenovic N., (1994), "Tabu search and simulated annealing in p-median problems", Talk at the Canadian Operational Research Society Conference, Montral.
- [98] Moritz, P. , Nishihara, R. , and Jordan, M. (2016), "A linearly-convergent stochastic L-BFGS algorithm", **In Artificial Intelligence and Statistics**, pp 249-258.
- [99] Morris, J.G. (1981), "Convergence of the Weiszfeld algorithm for Weber problems using a generalized distance function", **Oper. Res.**, 29, pp 37-48.
- [100] Morris, J.G., Verdini, W.A. (1979), "A simple iterative scheme for solving minisum facility location problems involving lp distances", **Oper. Res.**, 27, pp 1180-1188.
- [101] Nazari, M., Fathali, J. (2018), "Reverse backup 2-median problem with variable coordinate of vertices", **Journal of Operational Research and Its Applications**, 15 (2), pp 63-88.
- [102] Nesterov, Y, and Polyak, B. T. (2006), "Cubic regularization of newton method and its global performance", **Mathematical Programming**, 108(1), pp 177-205.
- [103] Nesterov, Y. (1983), "A method of solving a convex programming problem with convergence rate $O(1/k^2)$ ", **In Doklady AN SSSR (translated as Soviet Mathematics Doklady)**, volume 269, pp 543-547.
- [104] Nesterov, Y. (2004), "Introductory Lectures on Convex Optimization: A Basic Course", **Springer**.

- [105] Nocedal, J. and Wright, S.(2006), "**Numerical Optimization**", Springer Science & Business Media.
- [106] Notash, M., Zandieh, M., Dorri Nokorani, B. (2015), "Using a Genetic Algorithm Approach for Designing Multi-objective Supply Chain Network", **IQBQ**, 18 (4), pp 183-203 .
- [107] Oren, S.S. and Luenberger, D.G. (1973), "Self-scaling variable metric (SSVM) algorithms. I. Criteria and sufficient conditions for scaling a class of algorithms," **Management Sci.** 20(5) pp 845-862.
- [108] Oren, S.S. and Spedicato, E.(1976), "Optimal conditioning of self-scaling variable metric algorithms," **Math. Programming** 10(1), pp 70-90.
- [109] Owen S.H., Daskin M.S., (1998), "Strategic facility location: A review", **EJOR**, 111, pp 423-447.
- [110] Pereira, M. A., Coelho L. C., Lorena L. A. N., and de L. C., (2015), "A hybrid method for the Probabilistic Maximal Covering Location-Allocation Problem", **Computers and Operations Research**, 57, pp 51-59.
- [111] Poon, C., LIANG, J. and SCHOENLIEB, C.(2018), "Local convergence properties of SAGA/Prox-SVRG and acceleration", Proceedings of the 35th Int. Conf. on Mach. Learn., 80, 4124–4132.
- [112] Qin Z., and Gao Y., (2017), "Uncapacitated p-hub location problem with fixed costs and uncertain flows", **Journal of Intelligent Manufacturing**, 28 (3), pp 705 - 716 .
- [113] Rahmani A., and MirHassani S.A., (2014), "A hybrid Firefly-Genetic Algorithm for the capacitated facility location problem", **Information Sciences**, 283, pp 70 - 78 .
- [114] Robbins, H. and Monro, S. , (1951), "A stochastic approximation method", **The Annals of Mathematical Statistics**, pp 400-407.
- [115] Rolland E., Schilling D.A., Current J. R., (1997), "An efficient tabu search heuristic for the p-median problem", **EJOR**, 96, pp 329-342.

- [116] Roux, N. L. , Schmidt, M. , and Bach, F. R. , (2012), "A stochastic gradient method with an exponential convergence rate for finite training sets", **In Advances in Neural Information Processing Systems**, pp 2663-2671.
- [117] Sahin G., and Sural H., (2007), "A review of hierarchical facility location models", **Computers and Operations Research**, 34 (8), pp 2310 - 2331.
- [118] Scaparra M.P., Scutella M. G., (2001), "Facilities, locations, customers: Building blocks of location models: A survey", **Technical Report: tr-01-18**, University of Piza, Italy.
- [119] Schraudolph, N. , Jin Yu, and Simon G"unter.(2007), "A stochastic quasi-Newton method for online convex optimization",**In Artificial Intelligence and Statistics**, pp 436-443.
- [120] Senne E. L. F., Lorena L. A. N., and Pereira M. A., (2005), "A branch-and-price approach top - median location problems", **Computers and Operations Research**, 32, pp 1655 - 1664 .
- [121] Shalev-Shwartz, S. and Zhang, T. (2013), "Stochastic dual coordinate ascent methods for regularized loss minimization", **Journal of Machine Learning Research**, 14, pp 567-599.
- [122] Shalev-Shwartz, S. and Zhang, T. (2016), "Accelerated proximal stochastic dual coordinate ascent for regularized loss minimization", **Mathematical Programming**, 155(1-2), pp 105-145.
- [123] Shamir, O. , Srebro, N. , and Zhang, T. (2014), "Communication-efficient distributed optimization using an approximate Newton-type method",**In International Conference on Machine Learning**, pp 1000-1008.
- [124] Slater P. J., (1982), "Locating central path in a graph", **Transportation Science**, 16, pp 1-18.
- [125] Soleimani A., Fathali J., Nazari M. (2019). "Single facility goal location problems with L_p norm", **Modern Research in Decision Making**, 3, pp 125–152.
- [126] Tamir A., (1996), "An $O(p \sqrt{n})$ algorithm for the p-median and related problems on tree graphs", **Oper. Res Letters**, 19, pp 59-64.

-
- [127] Tansel B. C., Francis R.L., Lowe T. J., (1983), "Location on networks: A survey", **Management Sci.**, 29, pp 482-511.
- [128] Teitz M. B., and Bart P., (1968), "Heuristic methods for estimating the generalized vertex median of a weighted graph", **Operations Research**, 16, pp 955 - 961.
- [129] Teye C., HBell M. G., and JBliemer M. C., (2017), "Urban intermodal terminals: The entropy maximising facility location problem", **Transportation Research Part B: Methodological**, 100, pp 64 - 81 .
- [130] Trinh, M. H., Lee, B.H., and Ahn, H.S. (2015), "The Fermat-Weber location problem in single integrator dynamics using only local bearing angles", **Auto Matica**, 59, pp 90-96.
- [131] Tripuraneni N, Stern M, Jin C, Regier J, Jordan MI (2018), "Stochastic cubic regularization for fast nonconvex optimization", **Advances in Neural Information Processing**, vol. 32 (Curran Associates, Red Hook, NY), pp 2899–2908.
- [132] Uno T. , and Katagiri H., (2008), "Single- and multi-objective defensive location problems on a network", **European Journal of Operational Research**, 188 (1), pp 76-84.
- [133] Vardi, Y., Zhang, CH. (2001), "A modified Weiszfeld algorithm for the Fermat-Weber location problem", **Math. Program.**, Ser. A 90, pp 559–566.
- [134] Varian, H. R. , (1975), "A Bayesian approach to real estate assessment". In: **S. E. Fienberg, A. Zellner, (eds.), Studies in Bayesian Econometrics and Statistics in Honour of Leonard J. Savage**, North-Holland, Amesterdam, pp 195-208.
- [135] Ward J.E., Wendell R.E., (1980), "A new norm for measuring distance which yields linear location problems", **Oper. Res**, 28, pp 836-844.
- [136] Ward J.E., Wendell R.E., (1985)," Using block norms for location modelling", **Operations Research**, 33, pp 1074-1090.
- [137] Weber, A. (1929), "Uber den standort der industrient", Tübingen, (1909), English Trans.: Theory of location of industries, C.J., Friedrich, ed., and Trans., **Chicago University Press, Chicago, Illinois**.

- [138] Weiszfeld, E. (1937), "Sur le point par lequel la somme des distances de n points donns est minimum", **Tohoku Math.**, 43, pp 355–386.
- [139] Wesolowsky G. O., (1974), "Location of the median line for weighted points", **Environment and planning**, 7, pp 163-170.
- [140] Wesolowsky, G. O. and Love, R. F., (1971)," Location of facilities with rectangular distance among point and area destinations",**Naval Research Logistics Quarterly**, 1, pp 83 - 90 .
- [141] Xu, P. , Roosta-Khorasani, F. , and Mahoney, M. W. (2019), "Newton-type methods for nonconvex optimization under inexact Hessian information", **Mathematical Programming**. doi:10.1007/s10107-019-01405-z
- [142] Xu, P. , Roosta-Khorasani, F. , and Mahoney, M. W. (2020), "Second-order optimization for non-convex machine learning: An empirical study",**Proceedings of the 2020 SIAM International Conference on Data Mining**.
- [143] Yuan G., Zengxin W. (2010). "Convergence analysis of a modified BFGS method on convex minimizations", **Computational Optimization and Applications**, 47, pp 237– 255.
- [144] Zellner, A. , Geisel, M. S. , (1968), "Sensitivity of Control to Uncertainty and from of the Criterion function",in **D. G. Watts(Ed.), The Future of Statistics**, Academic Press, Inc., NewYork, pp 269-289.
- [145] Zhang J., (2006), "Approximating the two-level facility location problem via a quasi-greedy approach", **Mathematical Programming**, 108 (1), pp 159 - 176 .
- [146] Zhou, D. , Xu, P. and Quanquan Gu. (2018), "Stochastic Variance-Reduced Cubic Regularized Newton Method",In **ICML**.

واژه‌نامه فارسی به انگلیسی

Heuristic	ابتکاری
Extension principal	اصل گسترش
Gradient projection algorithm	الگوریتم تصویر گرادیان
Curvature	انحناء
Critical	بحرانی
Unweighted	بدون وزن
Linear programming	برنامه ریزی خطی
Infimum	بزرگ ترین کران پایین
Continuously differentiable	به طور پیوسته مشتق پذیر
Optimal	بهینه
Optimization	بهینه سازی
Maximum	بیشینه
Complexity	پیچیدگی
loss function	تابع زیان
Convex function	تابع محدب
Characteristic function	تابع مشخصه
Concave function	تابع مقعر
Objective function	تابع هدف
Convexity	تحدب
Generalized	تعمیم یافته – عمومیت داده شده
Extended	تعمیم یافته – گسترش یافته
Concavity	تقعر
Iteration	تکرار
Single facility	تک وسیله ای
Sparse	تنک

Develop	توسعه دادن
Search	جستجو
Direction	جهت
Optimal solution	جواب بهینه
Feasible solution	جواب شدنی
Unique solution	جواب یکتا
Polynomial	چند جمله‌ای
Optimism	خوش بینانه
Precision	دقت
Vertex	راس
Coordinate descent method	روش مختصات کاهشی
Interior point method	روش نقطه درونی
Hypomonotone	زیریکنوا
Consistence	سازگار
Accurate	صحیح – دقیق
Increasing	صعودی
Right-hand-side coefficients	ضرایب سمت راست
Constraints coefficients	ضرایب محدودیت ها
Cost coefficients	ضرایب هزینه
Lagrangian multiplier	ضریب لاگرانژ
Uncertainty	عدم قطعیت
Imprecise	غیردقیق
Distance	فاصله
Least squares	کم‌ترین مربعات
Supremum	کوچک‌ترین کران بالا
Newton-like step	گام شبه نیوتن
Gradient	گرادیان
Lipschitz	لیپ‌شیتس
Jacobian matrix	ماتریس ژاکوبی
Symmetric	متقارن
Feasible set	مجموعه شدنی
Convex set	مجموعه محدب

Concave set	مجموعه مقعر
Strictly convex	محدب اکید
Constraint	محدودیت
Center	مرکز
Regularized optimization problems	مسائل بهینه‌سازی تنظیم‌شده
Customer	مشتری
Directional derivative	مشتق جهتی
Inverse	معکوس
Location	مکان‌یابی
Median	میانه
Infinite	نامتناهی
Frobenius norm	نرم فروبینیوس
Nuclear norm	نرم هسته‌ای
Norm	نرم
Normality	نرمال بودن
Decreasing	نزولی
Demand point	نقطه تقاضا
Crossover point	نقطه گذر
Core	هسته
Smooth	هموار
Weight	وزن
Weighted	وزندار
Edge	یال

واژه‌نامه انگلیسی به فارسی

Accurate	صحيح - دقيق
Armijo	آرمیژو
Characteristic function	تابع مشخصه
Complexity	پیچیدگی
Concave function	تابع مقعر
Concave set	مجموعه مقعر
Concave	مقعر
Concavity	تقعر
condition	شرط
Consistence	سازگار
Constraint	محدودیت
Constraints coefficients	ضرایب محدودیت ها
Continuously differentiable	به طور پیوسته مشتق پذیر
Contradict	متناقض
Convergence order	مرتبه همگرایی
Convex function	تابع محدب
Convexity	تحدب
Coordinate descent method	روش مختصات کاهشی
Cost coefficients	ضرایب هزینه
Cost	هزینه
Curvature	انحناء
Customer	مشتری
Decision maker	تصمیم گیرنده
Decision variable	متغیر تصمیم گیری
Decreasing	نزولی

Demand point	نقطه تقاضا
Descent direction	جهت کاهشی
Develop	توسعه دادن
Directional derivative	مشتق جهتی
Distance	فاصله
Edge	یال
Eigen value	مقدار ویژه
error	خطا
Euclidean	اقلیدسی
Extended	تعمیم یافته – گسترش یافته
Feasible solution	جواب شدنی
Finite	متناهی – محدود
First-order necessary condition	شرط لازم مرتبه اول
Frobenius norm	نرم فروبینیوس
Generalized	تعمیم یافته – عمومیت داده شده
Global solution	جواب سراسری
Goldstein	گلدشتاین
Gradient	گرادیان
Hessian	هسی
Heuristic	ابتکاری
Ill condition	بد حالت
Imprecise	غیردقیق
Infimum	بزرگ‌ترین کران پایین
Infinite	نامتناهی
Inverse	معکوس
Iteration	تکرار
Jacobian matrix	ماتریس ژاکوبی
Lagrangian multiplier	ضریب لاگرانژ
Large scale	مقیاس بزرگ
Least squares	کم‌ترین مربعات
Line search	جستجوی خطی
Linear programming	برنامه ریزی خطی

Lipschitz	لیپ‌شیتس
Local	موضعی
Location	مکان‌یابی
loss function	تابع زیان
matrix	ماتریس
Maximum	بیشینه
minimizer	کمینه‌کننده
Newton-like step	گام شبه نیوتن
Node	گره
Norm	نرم
Nuclear norm	نرم هسته‌ای
Objective function	تابع هدف
Optimal solution	جواب بهینه
Optimal	بهینه
Optimism	خوش بینانه
Optimization	بهینه‌سازی
Pessimistic	بدبینانه
Positive definite	معین مثبت
Positive semidefinite	نیمه معین مثبت
Precision	دقت
Quasi newton	شبه نیوتن
Regularized optimization problems	مسائل بهینه‌سازی تنظیم‌شده
Sample	نمونه
Secant equation	معادله سکانت
Second-order necessary condition	شرط لازم مرتبه دوم
Second-order sufficient condition	شرط کافی مرتبه دوم
Single facility	تک‌وسيله‌ای
Singular value decomposition	تجزیه مقدار تکین
Smooth	هموار
Sparse	تنک
Stability	پایداری
Steepest descent	تندترین کاهش

Step size	طول گام
Stochastic gradient	گرادیان تصادفی
Strict global	سراسری اکید
Strict local	موضعی اکید
Strictly convex	محدب اکید
Strong Wolfe	ولف قوی
Strongly convex	محدب قوی
Supremum	کوچک‌ترین کران بالا
Symmetric	متقارن
Total	کلی
Uncertainty	عدم قطعیت
Unique solution	جواب یکتا
Unweighted	بدون وزن
Vertex	راس
Weighted	وزندار
Search	جستجو
Global	سراسری

Aabstract

Location theory is one of the most important topics in optimization and operations research. In location problems, the goal is to find the location of one or more facilities such that some criteria such as transportation costs, customer distance traveled, total service time, and cost of servicing are optimized.

In this thesis, we consider the ideal location problems in which the location of a number of customers on a plane is given, and the ideal is locating the facility in the distance R_i , from the i -th client. However, in most instances, the solution of this problem doesn't exist. Therefore, the minimizing sum of errors is considered. For this purpose, we consider two models of symmetric and asymmetric loss (error) function and use the distance function ℓ_p norm to define these loss functions. In order to minimize the symmetric loss function, we use the Weiszfeld like method, Gauss-Newton and Imperialist competitive algorithm (ICA) and compare the numerical results obtained by solving these algorithms.

In addition, in another part of this thesis, we consider the asymmetric loss function and explain its interesting features. Using the some modified BFGS methods, Imperialist competitive algorithms, and the Weiszfeld-like algorithm, we minimize this loss function and present the results of the implementation of these algorithms and compare these results with each other.

Also, in the final part of this thesis, we seek to solve large-scale problems using stochastic optimization methods. Stochastic optimization methods perform very well in large-scale problems by reducing computational cost and achieve less objective function value in less time, which has led to the popularity of these methods among researchers in the field of optimization. We have also solved the problem of ideal location under the symmetric loss function by using Stochastic methods and have obtained very good results, which we have presented and described in the final part of the thesis using various figures and tables.

Keywords: Ideal location problem, loss function, Imperialist competitive algorithm, Wiesfeld-like method, BFGS methods, Stochastic optimization methods.



Faculty Of Mathematical Sciences

PhD Thesis in Operations research

Methods for solving compromised location problems

By: Aria Soleimani Koorandeh

Supervisor:

Jafar Fathali

Advisor:

Ahmad Nezakati

September 2021