



دانشگاه صنعتی شریف
دانشکده علوم ریاضی

عنوان:

حل مسئله جانمایی کره‌ها با روش‌های کاهش گرادیان و شبه نیوتونی

Solving Sphere Packing Problem using Gradient
Descent and Quasi-Newton Methods

اعضای گروه

محمد محسن عباس زاده
سینا دانشگر

نام درس

آشنایی با بهینه‌سازی

نیم‌سال اول ۱۴۰۳-۱۴۰۴

نام استاد درس

کسری علیشاهی

چکیده:

این مطالعه به بررسی مسئله‌ی جانمایی بهینه کره‌ها در یک محیط محصور با استفاده از روش‌های کاهش گرادیان و شبه نیوتونی پرداخته است. در این تحقیق، دو روش بهینه‌سازی مستقیم مکان کره‌ها با استفاده از تابع انرژی پتانسیل و بهینه‌سازی تکراری مکان کره‌ها و شعاع ظرف مورد بررسی قرار گرفته‌اند. نتایج نشان می‌دهد که روش‌های مختلف در شرایط متفاوت می‌توانند کارایی و دقت‌های متفاوتی ارائه دهند. به علاوه، مسئله دوگان، کاربردهای علمی و صنعتی این مسئله و راهکارهای آتی برای بهبود نتایج بررسی شده‌اند.

واژه‌های کلیدی: تابع انرژی، شبه نیوتون، کاهش گرادیان، چیدمان بهینه کره‌ها

۱ مقدمه

بسته‌بندی بهینه‌ی کره‌ها در یک محیط محصور، یکی از مسائل پایه‌ای در ریاضیات و علوم مهندسی است. این مسئله ارتباط مستقیمی با بهینه‌سازی ترکیبیاتی، نظریه‌ی کدگذاری، و مکانیک مواد دارد. در این مطالعه، به بررسی روش‌های مختلف حل این مسئله پرداخته و عملکرد آن‌ها را از نظر کارایی و دقت مورد تحلیل قرار داده‌ایم.

۱.۱ تعریف مسئله

مسئله‌ی مورد بررسی، یافتن کمترین شعاع R برای یک کره‌ی بزرگتر است که بتواند n کره‌ی کوچکتر با شعاع واحد را در خود جای دهد، به طوری که هیچ‌کدام از کره‌های کوچکتر با یکدیگر تداخل نداشته باشند. این مسئله معادل است با یافتن بهینه‌ترین چیدمان کره‌های کوچکتر درون کره‌ی بزرگتر.

۲ روش‌های مورد استفاده

در این تحقیق، دو روش بهینه‌سازی برای حل مسئله بررسی شده‌اند:

۱.۲ روش کاهش گرادیان

این روش یکی از تکنیک‌های رایج در بهینه‌سازی عددی است که از مشتقات مرتبه اول تابع هدف برای یافتن جهت نزول استفاده می‌کند. در این روش، با حرکت در جهت منفی گرادیان، مقدار تابع هدف کاهش می‌یابد. یکی از چالش‌های این روش انتخاب گام مناسب برای به‌روزرسانی موقعیت کره‌ها در فضای جستجو است.

۲.۲ روش شبه نیوتون

این روش یک تکنیک پیشرفته‌تر نسبت به کاهش گرادیان است که از تقریب‌های ماتریس هسیان برای به‌روزرسانی گرادیان استفاده می‌کند. در این روش، اطلاعات مشتق دوم به‌طور ضمنی تخمین زده شده و سرعت همگرایی بهبود می‌یابد. یکی از مزایای این روش نسبت به کاهش گرادیان، نرخ همگرایی سریع‌تر آن به مینیمم محلی است.

۳ ایده‌های حل مسئله

۱.۳ بهینه‌سازی مستقیم مکان کره‌ها با تعریف تابع انرژی پتانسیل

• ایده اصلی:

در این روش، مسئله به عنوان یک مسئله بهینه‌سازی سراسری در نظر گرفته می‌شود. هدف، یافتن مکان‌هایی برای کره‌ها است که یک تابع انرژی پتانسیل مناسب را کمینه کند.

تابع پتانسیل زیر را در نظر بگیرید:

$$\phi(x_1, x_2, \dots, x_n) = \sum_{i \neq j} \frac{1}{\|x_i - x_j\|} + \sum_i \|x_i\|^2$$

که در آن $x_1, x_2, \dots, x_n \in \mathbb{R}^3$ نشان دهنده مکان n نقطه است و $\| \cdot \|$ نرم ۲ می‌باشد.

هدف کمینه کردن این تابع است.

جمله اول، باعث می‌شود که نقاط تا جای ممکن از هم دور شوند. جمله دوم، نقاط را در یک محدوده در نزدیکی مبدأ قرار می‌دهد.

این باید باعث شود تا نقاط درون یک کره، به صورت مناسب توزیع شوند.

- تخمین شعاع کره بزرگتر پس از یافتن مکان‌های بهینه برای کره‌ها با استفاده از روش کاهش گرادیان، فاصله نزدیک‌ترین نقاط را محاسبه کرده و بر ۲ تقسیم می‌کنیم. حاصل را r در نظر می‌گیریم. سپس فاصله دورترین نقطه از مبدا را با r جمع کرده (عدم همپوشانی کره بزرگتر با کره‌های لب مرز) و نهایتاً با تقسیم بر r شعاع کوچکترین کره‌ای که بتوان n کره واحد را در آن جای داد، به دست می‌آید.

- مشکلات و تلاش‌ها برای بهبود نتایج

یکی از مشکلات اصلی این روش، توزیع غیریکنواخت نقاط است. به دلیل ساختار تابع انرژی پتانسیل، نقاط در نزدیکی مبدأ فشرده‌تر هستند و هرچه از مبدأ دورتر می‌شویم، فشرده‌گی آن‌ها کمتر می‌شود. این مسئله، باعث می‌شود که با بزرگ شدن تعداد کره‌های کوچک، جواب نهایی با خطا همراه باشد. مشکل دیگر، افتادن در بهینه‌های محلی است که ناشی از انتخاب طول گام با روش Backtracking است. برای رفع این مشکل، تلاش‌هایی برای بهینه‌سازی تابع پتانسیل انجام دادیم، اما با شکست مواجه شد.

– **تغییر ضرایب و توان‌ها:** یکی از تلاش‌ها، تغییر ضرایب و توان‌های جملات اول و دوم تابع پتانسیل بود. هدف این بود که با تغییر قدرت دافعه و جاذبه، توزیع یکنواخت‌تری از نقاط ایجاد شود. اما این تغییرات، نه تنها باعث بهبود نتایج نشد، بلکه در بسیاری از موارد، خطا را افزایش داد.

– **اضافه کردن جمله سوم:** تلاش دیگری که انجام شد، اضافه کردن یک جمله‌ی سوم به تابع پتانسیل بود. ایده‌ی اصلی این بود که یک پوسته‌ی کروی خارجی با بار همنام با نقاط فرض شود که نقاط دور از مرکز را به سمت مرکز هل دهد. اما این تلاش نیز با شکست مواجه شد و نتایج بهتری حاصل نشد.

– **استفاده از روش نیوتون:** برای بهبود دقت نتایج، از روش نیوتون استفاده شد. از آنجا که روش نیوتون به نقطه شروع حساس است، جواب نهایی روش کاهش گرادیان به عنوان نقطه شروع روش نیوتون در نظر گرفته شد. اما این کار، تنها باعث افزایش زمان اجرا شد و بهبود قابل توجهی در دقت نتایج حاصل نشد. این مسئله احتمالاً به دلیل گیر افتادن در یک مینیمم محلی بوده است.

– **طول گام ثابت:** در نهایت، به جای استفاده از طول گام متغیر که با روش Backtracking به دست می‌آمد، از یک طول گام ثابت $\alpha = 0.2$ استفاده شد. طول گام‌های مختلفی امتحان شد که در نهایت این مقدار، جواب‌های بهتری داد. این کار، باعث بهبود نسبی نتایج و کاهش زمان اجرا شد.

اما همچنان مشکل اصلی، یعنی توزیع غیریکنواخت نقاط، برطرف نشد...

۲.۳ بهینه‌سازی تکراری مکان کره‌ها و شعاع ظرف

- ایده اصلی: در این روش، به جای بهینه‌سازی مستقیم مکان کره‌ها، یک رویکرد تکراری برای یافتن چیدمان بهینه کره‌ها و شعاع کره‌ی بزرگتر ارائه می‌شود. در این روش، ابتدا مکان‌های تصادفی برای کره‌ها در یک کره‌ی بزرگتر با شعاع اولیه‌ی R تولید می‌شود. سپس، به صورت تکراری، مکان کره‌ها و شعاع کره‌ی بزرگتر بهینه می‌شوند.

۱.۲.۳ تابع انرژی پتانسیل

تابع انرژی پتانسیل در این روش، مشابه روش اول است و از دو بخش انرژی دافعه و انرژی محدودسازی تشکیل شده است:

$$O_{ij} = \max(0, 2 - \|x_i - x_j\|)$$

$$O_{i*} = \max(0, \|x_i\| + 1 - R)$$

$$E(x) = \sum_{i=1}^n \sum_{i \neq j} O_{ij}^x [i < j] + \sum_{i=1}^n O_{i*}^x$$

با این تفاوت که در این روش، شعاع R نیز به عنوان یک متغیر بهینه‌سازی در نظر گرفته می‌شود. توضیح: جمله اول، مشابه روش اول، به دنبال دور کردن نقاط از یکدیگر است. جمله دوم، نقاط را درون کره بزرگتر محدود می‌کند. اگر نقطه‌ای از مرز کره (با در نظر گرفتن شعاع ۱ برای کره کوچک) خارج شود، یک نیروی بازگرداننده به آن اعمال می‌شود.

۲.۲.۳ الگوریتم BFGS

الگوریتم BFGS یک روش شبه‌نیوتنی برای بهینه‌سازی توابع غیرخطی است. این الگوریتم از یک تقریب از ماتریس هسین استفاده می‌کند تا جهت جستجو را تعیین کند. در مقایسه با روش گرادیان کاهشی، BFGS معمولاً سریع‌تر به حل همگرا می‌شود زیرا از اطلاعات انحنا تابع هدف استفاده می‌کند.

— مقداردهی اولیه:

۱. موقعیت‌های اولیه به صورت تصادفی مقداردهی می‌شوند.
۲. تقریب ماتریس هسین با یک ماتریس همانی مقداردهی می‌شود.
۳. گرادیان اولیه محاسبه می‌شود.

– تکرار:

۱. تعیین جهت جستجو: جهت با استفاده از فرمول $direction = -H * grad_{old}$ محاسبه می‌شود. این جهت، جهتی است که انتظار می‌رود تابع هدف در آن جهت کاهش یابد.
۲. جستجوی خطی: اندازه گام با استفاده از جستجوی خطی تعیین می‌شود. جستجوی خطی به دنبال یافتن بهترین اندازه گام در جهت جستجو است، به طوری که تابع هدف به میزان کافی کاهش یابد.
۳. موقعیت‌ها با استفاده از فرمول $positions = positions + \alpha * direction$ به روزرسانی می‌شوند.
۴. گرادیان جدید در موقعیت‌های جدید محاسبه می‌شود.
۵. تقریب ماتریس هسین با استفاده از فرمول BFGS به‌روزرسانی می‌شود. این فرمول از تغییرات در گرادیان و موقعیت‌ها برای بهبود تقریب ماتریس هسین استفاده می‌کند.
۶. اگر تغییرات در موقعیت‌ها، کمتر از یک آستانه تحمل (tolerance) باشد، الگوریتم متوقف می‌شود.

– پارامترها و نحوه تنظیم آن

۱. حداکثر تعداد تکرار: این پارامتر تعیین می‌کند که الگوریتم حداکثر چندبار تکرار شود. مقدار بزرگتر به الگوریتم فرصت بیشتری برای همگرایی می‌دهد. اما زمان اجرا افزایش می‌یابد. مقدار کوچکتر ممکن است باعث شود که الگوریتم به حل بهینه نرسد. معمولاً مقادیر بین ۱۰۰ تا ۱۰۰۰ مناسب هستند.
۲. آستانه تحمل: این پارامتر تعیین می‌کند که تغییرات در موقعیت‌ها باید کمتر از چه مقداری باشد تا الگوریتم متوقف شود. مقدار کوچکتر، دقت بیشتری را به همراه می‌آورد. اما زمان اجرا را افزایش می‌دهد.
۳. طول گام: این پارامتر به صورت خودکار توسط جستجوی خطی تنظیم می‌شود.
۴. کاهنده طول گام: این پارامتر تعیین می‌کند که اندازه گام در هر تکرار جستجوی خطی، چقدر کاهش یابد. مقدار کوچکتر، باعث می‌شود تا جستجوی خطی دقیق‌تر باشد. اما زمان اجرا را افزایش می‌دهد. معمولاً مقدار ۰/۵ مناسب است.

۵. ثابت کوچک: این پارامتر در شرط کافی برای کاهش تابع هدف استفاده می شود. معمولاً مقدار ۰/۰۱ مناسب است.

۴ بررسی نتایج، مقایسه جواب‌ها در دو روش، و مقایسه با جواب بهینه

در شکل ۱ مقایسه شعاع‌های دو روش با شعاع بهینه آورده شده است.

شعاع بهینه (تئوری/تجربی)	تابع بهینه سازی متلب	روش دوم	روش اول	تعداد کره‌ها (n)
	شعاع تخمینی	شعاع تخمینی	شعاع تخمینی	
2.414213562	2.414217031	2.414219423	2.414213562	5
2.832464561	2.833973681	2.779494457	2.872505157	10
3.000000000	2.999999907	3.110159119	3.000000000	13
3.473538962	3.483743034	3.531900921	3.627878501	20
3.916491662	3.958442933	3.932238212	4.241683460	30
4.255332954	4.291424986	4.266693658	4.629004683	40
4.550454341	4.616370344	4.569101208	4.815877272	50
5.111741006	5.173238678	5.165086182	5.492972249	73
5.635980816	5.676452496	5.630891589	6.204832054	100
6.381471145	6.479237736	6.396721151	7.083588286	150

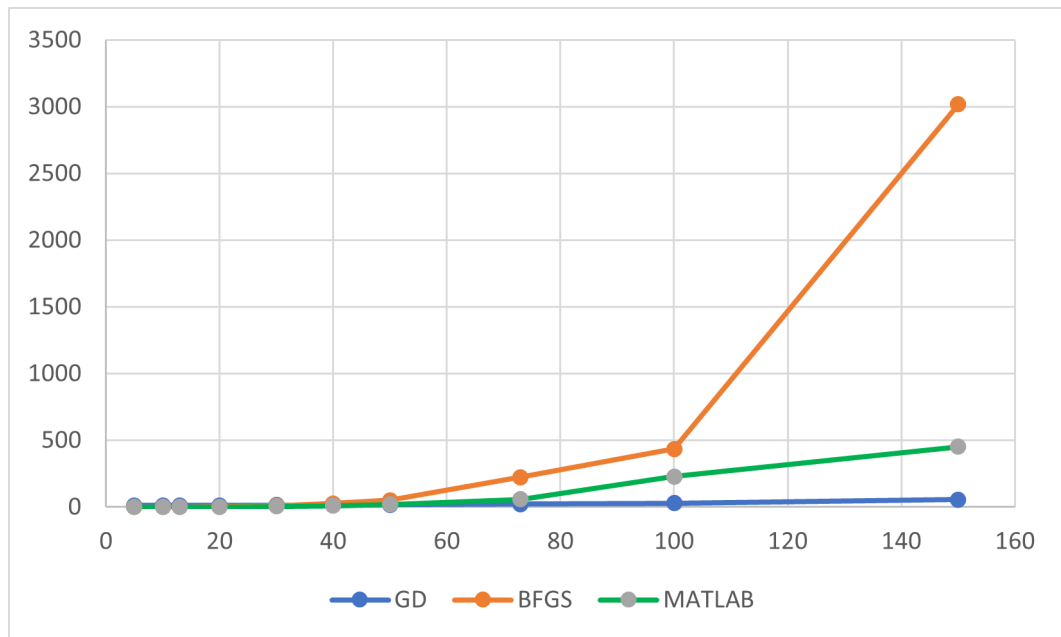
شکل ۱: مقایسه شعاع‌های دو روش، با شعاع بهینه

در شکل ۲ خطای دو روش و همچنین زمان اجرای الگوریتم‌های آنان ارائه شده است.

تابع بهینه سازی متلب		روش دوم		روش اول		تعداد کره‌ها (n)
زمان	درصد خطا	زمان	درصد خطا	زمان	درصد خطا	
0.518	0.00014367	0.518	0.0002427	12.378	0	5
0.543	0.053279393	0.654	1.8701065	12.434	1.413630967	10
0.637	3.0919E-06	0.947	3.6719706	13.17	0	13
1.266	0.29376587	2.098	1.6801872	13.525	4.443293734	20
4.087	1.071144147	10.191	0.4020575	14.550	8.303140318	30
9.953	0.848160016	25.601	0.2669757	14.863	8.781257159	40
20.427	1.448558723	51.996	0.4097803	17.303	5.832888567	50
56.852	1.203067047	222.587	1.0435813	21.718	7.457953020	73
227.767	0.718094702	435.574	0.0902989	28.423	10.09320749	100
451.455	1.532038436	3018.564	0.2389732	54.455	11.00243383	150

شکل ۲: مقایسه درصد خطا و زمان اجرای دو روش

همان‌طور که در شکل ۲ مشاهده می‌شود با افزایش تعداد کره‌ها خطای روش اول بیشتر می‌شود، اما روش دوم همواره درصد خطای قابل قبولی دارد. در عوض زمان اجرای الگوریتم دوم با افزایش ابعاد مساله به صورت نمایی زیادتر می‌شود. مشاهده‌ی دیگر این است که در اکثر موارد روش دوم نسبت به تابع بهینه‌سازی خود متلب کمی دقت بالاتری دارد، اما زمان اجرای آن هم با افزایش مقدار n رشد بیشتری دارد. شکل‌های ۳ و ۴ به ترتیب نمودار زمان حل و کیفیت جواب‌های هر الگوریتم را بر حسب تعداد کره‌های واحد نشان می‌دهد.



شکل ۳: زمان حل الگوریتم‌های استفاده شده

شکل ۵ نمایش سه بعدی جواب‌های به دست آمده برای مقادیر مختلف n است.

۵ مسئله دوگان

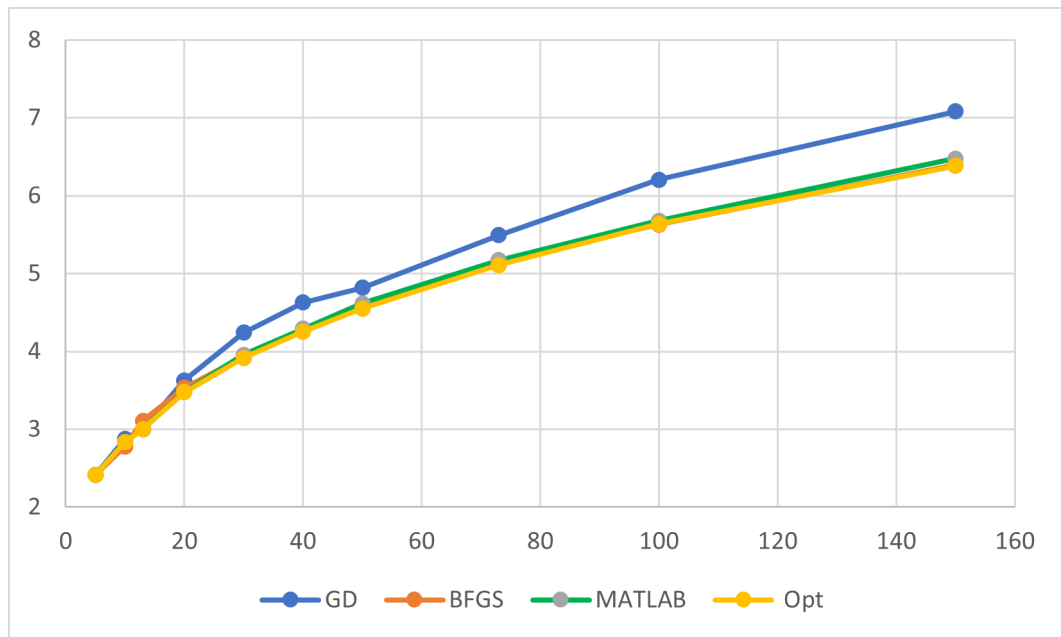
دوگان مسئله اصلی را، اینطور تعریف می‌کنیم که:

یک کره به شعاع R داریم، حداکثر چند گوی واحد می‌توانیم در داخل آن جای دهیم؟

به عبارتی، مسئله اصلی (Packing) به دنبال یافتن بهترین چیدمان برای تعداد مشخصی گوی در یک فضا است، در حالی که مسئله دوگان، به دنبال یافتن حداکثر تعداد گوی‌هایی است که می‌توان در یک فضای مشخص (مثلاً کره‌ای با شعاع R) جای داد.

همانطور که در مقدمه ذکر کردیم، این مسئله کاربردهای گسترده‌ای در علوم و مهندسی دارد، من جمله موارد زیر:

- شیمی: در بررسی نحوه قرارگیری مولکول‌ها در یک فضا، به خصوص در نانوشیمی و طراحی داروها.
- صنایع بسته‌بندی: بهینه‌سازی فضای مورد استفاده در بسته‌بندی محصولات
- فیزیک: بررسی ساختار اتمی مواد و آرایش ذرات در سیستم‌های مختلف



شکل ۴: کیفیت جواب‌های دو روش استفاده شده، نسبت به جواب متلب و جواب بهینه

۶ جمع‌بندی

به طور خلاصه، روش اول برای مسائلی که نیاز به دقت بالایی ندارند و سرعت اجرا در اولویت است مناسب است. در مقابل، روش دوم برای مسائلی که دقت بالا اهمیت بیشتری دارد و زمان اجرا محدودیتی ندارد، انتخاب بهتری است. همچنین، می‌توان از ترکیب این دو روش استفاده کرد.

۷ کارهای آتی

- گوی‌هایی با شعاع‌های متفاوت
- بررسی تاثیر شکل ظرف
- بسته‌بندی کره در ابعاد بالاتر
- استفاده از روش‌های یادگیری ماشین برای حل مسئله
- بررسی اثر نیروهای خارجی



$n = 50$



$n = 40$



$n = 30$



$n = 100$



$n = 73$

شکل ۵: نمایش سه بعدی جواب‌ها

۱.۷ سوال جالب پیشنهادی

- چگونه می‌توان از خواص خود - سازمان‌دهی سیستم‌های فیزیکی (مثل نانو ذره‌ها) برای ایجاد چیدمان بهینه کره‌ها استفاده کرد؟