بسم الله الرحمن الرحیم

دانشگاه گیلان

دانشکدهٔ علوم پایه
گروه ریاضی
(گرایش کاربردی)

روش نیوتن تعمیم یافته برای حل مسالهٔ برنامه ریزی خطی

از:

مریم پرندگان کاظمی پور

استاد راهنما:

دکتر سعید کتابچی

استاد مشاور:

دکتر مازیار صلاحی

شهریور ۱۳۸۷

تقدیم به

مولای متقیان امام علی (ع)

مدح علی (ع) چه جلوه نماید ز ما

او را چو بهتر از همه یزدان ستوده است

و

بهترین های زندگی ام

پدر و مادر عزیزم

تقدیر و تشکر

توفیق دادار بزرگ اگر نمی‌بود، این کار رنگ بودن نمی‌گرفت. او را به سبب این توفیق وهمهٔ توفیق‌ها شکر می‌نهم. بدون ابراز سپاس از تمام کسانی که مرا یاری نمودند نیز نمی‌توان بر این سطور نقطهٔ پایان گذاشت. از پدر و مادر عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان که در سردترین روزگاران بهترین پشتیبان است، سپاس بی‌کران دارم.

از استاد راهنمای گرامی، جناب آقای دکتر سعید کتابچی، که همواره با راهنمایی‌های ارزشمندشان مشوقم بودند تشکر و قدردانی می‌نمایم. همچنین از استاد مشاور ارجمند، جناب آقای دکتر مازیار صلاحی، که با زحمات بی‌دریغشان در تمام مراحل تدوین این پایان نامه راهگشای اینجانب بوده‌اند، کمال تشکر و سپاسگزاری را دارم. مراتب قدردانی خود را از داوران محترم این پایان نامه، جناب آقای دکتر کیانپور و جناب آقای دکتر امیر تیموری، و نمایندهٔ محترم تحصیلات تکمیلی، جناب آقای دکتر سهله، اعلام می‌کنم.

و در پایان از خواهر و برادر عزیزم و تمام دوستانم که یاورم بودند، سپاسگزارم.

فهرست مطالب

عنوان | صفحه

فهرست جدول‌ها

عنوان ... صفحه

فهرست شکل‌ها

چکیدهٔ فارسی:

روش نیوتن تعمیم یافته برای حل مسئله برنامه ریزی خطی
مریم پرندگان کاظمی‌پور

جواب مسئله برنامه ریزی خطی را می‌توان با مینیمم‌سازی تابع جریمه خارجی مسئله دوگان (اولیه) به‌دست آورد. این تابع درجه دوم و قطعه‌ای محدب و دیفرانسیل‌پذیر است. روش نیوتن تعمیم یافته برای مینیمم‌سازی این تابع پیشنهاد می‌شود و این روش در چند خط برنامه به زبان Matlab قابل ارائه است.

واژه‌های کلیدی: روش نیوتن، روش جریمه، کمترین نرم جواب، ماتریس هسین، هسین تعمیم یافته

Abstract:

The Generalized Newton Method for Linear Programming

M.Parandegan Kazemipour

The solution of the LP problem can be obtained by minimization of the exterior penalty function of the dual(primal) problem. This function is piecewise quadratic convex and diffrentiable. The generalized Newton method is proposed for minimizing of this function and this Newton method can be given in a few lines of Matlab code.

Key words: Newton method, penalty-method, minimum norm solution, hessian matrix, generalized hessian matrix.

مقدمه

برنامه ریزی خطی با مسئلهٔ مینیمم‌سازی یا ماکزیمم‌سازی یک تابع خطی با محدودیت‌های خطی در شکل مساوی و یا نامساوی سروکار دارد. مسئلهٔ برنامه ریزی خطی را ابتدا جرج بی. دانتزیک، در حدود سال ۱۹۴۷ ابداع کرد ولی کار او تا سال ۱۹۵۹ ناشناخته باقی ماند. امروزه برنامه ریزی خطی فایده و مزیت خود را به‌عنوان مدلی مهم برای بسیاری از مسائل تخصیص و پدیده‌های اقتصادی نشان داده است. گسترش دائماً فزایندهٔ کاربردهای آن نشان دهندهٔ اهمیت برنامه ریزی خطی برای فرمول‌بندی مسائل است.

در سال ۱۹۴۹ جرج بی دانتزیک ((روش سیمپلکس)) را برای حل برنامه‌های خطی ارائه کرد. از آن به بعد افراد زیادی روشهای زیادی را برای حل این مسائل معرفی کردند. از جمله روش نیوتن یکی از روش‌هایی است که برای حل مسئلهٔ برنامه ریزی خطی به کار گرفته می شود. C.Kanzo، L.Qi. و H.Qi. در سال ۲۰۰۰ به این روش پرداختند [٦]؛ اما نتایج محاسباتی ارائه نکردند. O.L. Mangasarian در سال ۲۰۰۲ روش نیوتن تعمیم یافته را برای حل دسته ای از مسائل برنامه ریزی خطی به کار گرفت وبرای اولین بار نتایج محاسباتی را برای دو رده از مسائل، Data mining و Machine learning، ارائه کرد [۸]. لازم به ذکر است که O.L. Mangasarian سالها قبل به این روش روی آورده بود. اما در سال ۲۰۰۲ به نتایج عددی دست یافت. A.I. Golikov، Yu.G. Evtushenko، N. Molaverdi، S. Ketabchi و در سال‌های ۲۰۰۴ و ۲۰۰۵ روش تابع لاگرانژ را برای حل دسته‌ای ازاین مسائل مطرح کردند[۴، ۳].

در این پایان نامه از روش نیوتن تعمیم یافته برای حل دو رده از مسائل برنامه‌ریزی خطی استفاده می‌کنیم. یک دسته مسائلی با محدودیت‌های زیاد(در حدود 10^6) و متغیرهای متوسط (10^2) هستند. روش براین حقیقت استوار است که تابع جریمهٔ مسئلهٔ برنامه ریزی خطی معادل مسئلهٔ کمترین نرم دوگان مسئله است. همگرایی این الگوریتم سراسری می‌باشد. یک ویژگی ساده این روش آن‌است که کد مطلب آن در یک برنامهٔ کوتاه ارائه می‌شود و نتایج محاسباتی نوید بخش هستند. دستهٔ دیگر مسائلی هستند که تعداد محدودیت‌ها کمتر از تعداد متغیرهاست. در چنین مسائلی با استفاده از تابع جریمهٔ مسئلهٔ دوگان جوابی برای مسئلهٔ اولیه می‌یابیم. اما در این حالت مشکلاتی وجود دارد و نتایج حاصل به خوبی نتایج ردهٔ اول نیستند.

در فصل اول مقدمات مورد نیاز جبر خطی و برنامه ریزی خطی و خلاصه‌ای از روش سیمپلکس آورده شده است.

و در بخش دوم مباحث بهینه‌سازی مرتبط با این پایان‌نامه آمده است. در این بخش روش نیوتن کلاسیک معرفی می‌شود که اساس روش نیوتن تعمیم یافته را تشکیل می‌دهد. در بخش سوم ماتریس هسین تعمیم یافته و شرایط وجود این ماتریس معرفی می‌شود. در فصل پایانی روش نیوتن تعمیم یافته برای دو رده از مسائل مطرح می‌گردد. در این فصل کدهای متلب استفاده شده موجود است و پس از جدول‌های محاسباتی نتیجه‌گیری و پیشنهاد برای ادامهٔ کار آورده شده است.

فصل اول

قضایا و مقدمات اولیه

۱—۱ بردارها

تعریف ۱—۱—۱ : فرض کنید V یک فضای برداری حقیقی باشد، یک ضرب داخلی روی V تابعی است که به هر جفت از بردارهای u,v از V ، عدد حقیقی $<u,v>$ را نسبت می‌دهد که ضرب داخلی این دو بردار نامیده می‌شود. هر گاه برای بردارهای w,v,u و اسکالرهای α و β داشته باشیم :

الف) $<u,v> \geq \circ$ و $<u,v> = \circ$ اگر و تنها اگر $u = \circ$.

ب) $<u,v> = <v,u>$.

ج) $<\alpha v + \beta u, w> = \alpha <v,w> + \beta <u,w>$.

مثال ۱

فرض کنید $x,y \in \mathbb{R}^n$ باشد ، در این صورت حاصلضرب داخلی آنها را به صورت زیر تعریف می‌کنیم:

$$<x,y> = x^T y = \sum_{i=1}^{n} x_i y_i$$

که $y = (y_1, y_2, \ldots, y_n)^T$ و $x = (x_1, x_2, \ldots, x_n)^T$.

قضیه ۱—۱—۲ (نامساوی کشی شوارتز[1]): اگر $y = (y_1, y_2, \ldots, y_n)^T$ و $x = (x_1, x_2, \ldots, x_n)^T$ آنگاه:

$$(\sum_{k=1}^{n} x_k y_k)^2 \leq (\sum_{k=1}^{n} x_k^2)(\sum_{k=1}^{n} y_k^2).$$

برهان : [۱۵] مراجعه شود.

تعریف ۱—۱—۳ : مجموعهٔ بردارهای $\{v_1, v_2, \ldots, v_n\}$ در R^n را بردارهای مستقل خطی گویند، هرگاه از $\sum_{i=1}^{n} c_i v_i = \circ$ بتوان نتیجه گرفت برای هر i ، $c_i = \circ$.

---

[1] Cauchy- schwartz

نرم بردارها

تعریف ۱—۱—۴ : فرض کنید x یک بردار n بعدی باشد. نرم بردار x که با $\|x\|$ نشان داده می‌شود، یک مقدار حقیقی است که توسط یک تابع پیوسته از $\mathbb{R}^n$ به R تعریف می‌شود که دارای خواص زیر است :

الف) برای هر بردار غیر صفر x داریم $\|x\| > 0$ و $\|x\| = 0$ اگر وفقط اگر $x = 0$.

ب) برای هر بردار دلخواه x و هر اسکالر α داشته باشیم $\|\alpha x\| = |\alpha|\|x\|$.

ج) برای هر بردار دلخواه x و y داشته باشیم $\|x + y\| \leq \|x\| + \|y\|$. این رابطه به نامساوی مثلثی مشهور است .

برای هر بردار دلخواه x و y روابط زیر برقرار است :

الف) $\|x\| = \|-x\|$.

ب) $|\|x\| - \|y\|| \leq \|x - y\|$.

تعریف ۱—۱—۵ : نرم۱ (نرم مجموع) برای هر $x \in \mathbb{R}^n$ به شکل زیر تعریف می‌شود:

$$\|x\|_1 = |x_1| + |x_2| + ... + |x_n| \quad .$$

تعریف ۱—۱—۶ : نرم۲ (نرم‌اقلیدسی) برای هر $x \in \mathbb{R}^n$ به صورت زیر است :

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

تعریف ۱—۱—۷ : نرم∞ (نرم‌ماکسیمم) برای هر $x \in \mathbb{R}^n$ به صورت زیر است :

$$\|x\|_\infty = \max_i |x_i| \quad .$$

تعریف ۱—۱—۸ : در حالت کلی اگر p یک عدد حقیقی بزرگتر یا مساوی ۱ باشد، آنگاه نرم p[2] به صورت زیر است:

$$\|x\|_p = (\|x_1\|^p + \|x_2\|^p + ... + \|x_n\|^p)^{1/p} \quad .$$

[2] Holder norm

مثال ۲

فرض کنید $x = (۱, ۱, -۲)^T$

$$\|x\|_۱ = |۱| + |۱| + |-۲| = ۴$$

$$\|x\|_۲ = \sqrt{۱^۲ + ۱^۲ + (-۲)^۲} = \sqrt{۶}$$

$$\|x\|_\infty = \max\{۱, ۱, ۲\} = ۲$$

تعریف ۱—۱—۹ : دو نرم $\|.\|_\mu$ و $\|.\|_\nu$ را معادل گویند، هرگاه اعداد مثبت α و β وجود داشته باشند به طوریکه:

$$\forall x \in R^n : \qquad \alpha\|x\|_\mu \leq \|x\|_\nu \leq \beta\|x\|_\mu \quad .$$

قضیه ۱—۱—۱۰ : برای نرم ۱ ، ۲ و∞ رابطه های زیر برقرار است :

الف) $\|x\|_۲ \leq \|x\|_۱ \leq \sqrt{n} \quad \|x\|_۲$.

ب) $\|x\|_\infty \leq \|x\|_۲ \leq \sqrt{n} \quad \|x\|_\infty$.

ج) $\|x\|_\infty \leq \|x\|_۱ \leq \sqrt{n} \quad \|x\|_\infty$.

برهان: به [۲] مراجعه شود.

۱—۲ ماتریسها

تعریف ۱—۲—۱ : تعداد سطرها (ستونهای) مستقل خطی ماتریس A را رتبهٔ ماتریس A گویند که با $rank(A)$ نشان می‌دهند.

تعریف ۱—۲—۲ : فرض کنید A ماتریس $m \times n$ باشد. A را یک ماتریس کامل گویند ، هرگاه $rank(A) = min\{m, n\}$.

تعریف ۱—۲—۳ : فرض کنید $\lambda \in \mathbb{C}$ موجود باشد به طوریکه برای ماتریس A و بردار $x \neq ۰$ داشته باشیم $Ax = \lambda x$ ، در این صورت λ را یک مقدار ویژهٔ A وبردار x را بردار ویژهٔ متناظر λ گویند.

تعریف ۱—۲—۴ : فضای حاصل از تمام ماتریسهای $m \times n$ را با $L(\mathbb{R}^m, \mathbb{R}^n)$ نمایش می‌دهیم .

ماتریسهای خاص

تعریف ۱—۲—۵ : فرض کنیدA ماتریس $m \times n$ باشد . ترانهادهٔ ماتریس A که با A^T نمایش داده می شود، ماتریسی $n \times m$ است که از تعویض سطرهای ماتریس A با ستونهای آن بدست می‌آید. به عبارت دیگر اگر $A = [a_{ij}]$ ، آنگاه $A^T = [a_{ji}]$.

تعریف ۱—۲—۶ : ماتریس مربعی A را قطری گویند اگر تمام مولفه های غیر از قطر اصلی آن صفر باشند. ماتریس قطری را معمولاً به صورت $D = diag(d_{۱۱}, d_{۲۲}, \ldots, d_{nn})$ نمایش می‌دهند که:

$$D = \begin{pmatrix} d_{۱۱} & ۰ & \ldots & ۰ \\ ۰ & d_{۲۲} & \ldots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \ldots & d_{nn} \end{pmatrix}.$$

تعریف ۱—۲—۷ : اگر همه درایه‌های روی قطر اصلی یک ماتریس قطری برابر ۱ باشد ماتریس را همانی گویند و با I نشان می‌دهند.

تعریف ۱—۲—۸ : ماتریس مربعی A را متقارن گویند هرگاه داشته باشیم $A^T = A$.

تعریف ۱—۲—۹ : فرض کنید A یک ماتریس $n \times n$ باشد . ماتریس B ، $m \times n$ ، را معکوس (وارون) ماتریس A گویند هرگاه $AB = BA = I$ و آن را با $A^{-۱}$ نشان می‌دهند.

تعریف ۱—۲—۱۰ : ماتریس متقارن A را معیّن مثبت گویند هرگاه $\forall x \neq ۰ : x^T Ax > ۰$. همچنین ماتریس متقارن A را نیمه معیّن مثبت گوییم هرگاه $\forall x \in \mathbb{R}^n : x^T Ax \geq ۰$.

نکته ۱—۲—۱۱ [۲]:

الف) اگر A معین مثبت باشد، $A^{-۱}$ نیز معین مثبت است .

ب) اگرA ماتریس $m \times n$ $(m \geq n)$ رتبه کامل باشد، $A^T A$ معین مثبت است .

مثال ۳

فرض کنید $A = \begin{pmatrix} ۲ & ۱ \\ ۱ & ۵ \end{pmatrix}$

$$x = \begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix} \neq ۰ \Rightarrow x^T A x = \begin{pmatrix} x_۱ & x_۲ \end{pmatrix} \begin{pmatrix} ۲ & ۱ \\ ۱ & ۵ \end{pmatrix} \begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix} = ۲(x_۱ + \frac{۱}{۲}x_۲)^۲ + \frac{۹}{۲}x_۲^۲$$

این رابطه نشان می‌دهد A معیّن مثبت است.

نرم ماتریسی

تعریف ۱—۲—۱۲ : فرض کنید A یک ماتریس $m \times n$ باشد. مشابه تعریف (۱—۱—۴)، نرم ماتریس A را که با $\|A\|$ نشان داده می‌شود، به عنوان تابعی از $\mathbb{R}^{m\times n}$ به $\mathbb{R}$ تعریف می‌کنیم هرگاه دارای خواص زیر باشد:

الف) $\|A\| > ۰$. $\|A\| = ۰$ اگر و فقط اگر A ماتریس صفر باشد.

ب) برای هر اسکالر α داشته باشیم $\|\alpha A\| = |\alpha|\|A\|$.

ج) برای ماتریس های A و B داشته باشیم $\|A + B\| \leq \|A\| + \|B\|$.

نکته ۱—۲—۱۳ : برای هر ماتریس مانند A، $\|A\| = \|A^T\|$.

قضیه ۱—۲—۱۴ : نرم ماتریسی $\|A\|_p$ براساس نرم برداری $\|.\|_p$ به صورت زیر تعریف می‌شود.:

$$\|A\|_p = \max_x \frac{\|Ax\|_p}{\|x\|_p} \quad .$$

برهان : به [۲] مراجعه شود.

در تعریف های زیر فرض کنید A ماتریس $m \times n$ باشد.

تعریف ۱—۲—۱۵ : نرم —۱ (نرم ستونی) به صورت زیر تعریف می‌شود:

$$\|A\|_۱ = \max_{۱\leq j\leq n} \sum_{i=۱}^{m} |a_{ij}| \quad .$$

تعریف ۱—۲—۱۶ : نرم —∞(نرم سطری) به صورت زیر تعریف می‌شود:

$$\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}| \quad .$$

تعریف ۱—۲—۱۷ : نرم — ۲ [3] به صورت زیر است :

$$\|A\|_2 = \sqrt{\lambda_{\max}} \ ,$$

که λ_{max} بزرگترین مقدار ویژهٔ ماتریس A^TA است .

تعریف ۱—۲—۱۸ : نرم فروبنیوس [4] به شکل زیر تعریف می‌شود:

$$\|A\|_F = [\sum_{j=1}^{n} \sum_{i=1}^{m} |a_{ij}|^2]^{1/2} \quad .$$

مثال ۴

فرض کنید $A = \begin{pmatrix} 5 & 2 \\ 3 & 1 \end{pmatrix}$

$$\|A\|_1 = \max\{1+2, 5+3\} = 8$$

$$\|A\|_\infty = \max\{2+5, 1+3\} = 7$$

$$AA^T = \begin{pmatrix} 5 & 13 \\ 13 & 34 \end{pmatrix}$$

مقادیر ویژه AA^T برابر ۳۸/۹۷۴۳ و ۰/۰۲۵۷ است .

$$\max\{38/9743, 0/0257\} = 38/9743 \quad , \quad \|A\|_2 = \sqrt{38/9743} = 6/2429$$

$$\|A\|_F = \sqrt{2^2 + 1^2 + 5^2 + 3^2} = 6/2449$$

نکته ۱—۲—۱۹ [۲]: برای هر ماتریس A، $\|A\|_1 = \|A^T\|_\infty$ و $\|A\|_\infty = \|A^T\|_1$.

قضیه ۱—۲—۲۰ : برای نرم ماتریسها روابط زیر برقرار است:

---

[3] spectral norm

[4] Frobenius