وزارت علوم تحقیقات و فناوری
دانشگاه تفرش
دانشکده علوم ریاضی

گروه ریاضی کاربردی

پایان نامه کارشناسی ارشد

ریاضی کاربردی، گرایش آنالیز عددی

یک روش گرادیان مزدوج جدید برای حل مسائل بهینه سازی نامقید

استاد راهنما:

دکتر محمد افضلی نژاد

استاد مشاور:

دکتر دوستعلی مژده

تدوین و نگارش:

محمد معلمی

شهریور ۱۳۸۹

وزارت علوم تحقیقات و فناوری
دانشگاه تفرش
دانشکده علوم ریاضی
گروه ریاضی کاربردی

پایان نامه کارشناسی ارشد

ریاضی کاربردی، گرایش آنالیز عددی

یک روش گرادیان مزدوج جدید برای حل مسائل بهینه سازی نامقید

استاد راهنما:

دکتر محمد افضلی نژاد

استاد مشاور:

دکتر دوستعلی مژده

تدوین و نگارش:

محمد معلمی

شهریور ۱۳۸۹

چکیده

حل مسئله‌ی مینیمم سازی نامقید $\min f(x)\ ,\ x\in \mathbb{R}^n$، که در آن $\mathbb{R}^n$ یک فضای اقلیدسی n بعدی و $f:\mathbb{R}^n \to \mathbb{R}$ تابعی به طور پیوسته مشتق پذیر است، را در نظر می گیریم. روش گرادیان مزدوج، دیدگاهی مفید و قوی برای حل مسائل بهینه سازی در مقیاس بزرگ است. در این پایان نامه روش گرادیان مزدوج لیو و استوری را که عملکرد عددی خوبی دارد، تحت یک جستجوی خطی جدید آرمیجو-گونه برای مینیمم سازی توابعی که مشتقات جزئی پیوسته دارند، مورد بررسی قرار می‌دهیم. به وسیله‌ی تخمین ثابت لیپ‌شیتز برای مشتق تابع هدف، طول گام مناسب در هر تکرار به دست می‌آید، که همگرایی سراسری را تضمین می کند و به بهبود کارایی روش گرادیان مزدوج لیو و استوری در محاسبات عملی می‌انجامد. سپس الگوریتم روش پیشنهادی را در محیط نرم افزاری ***MATLAB ۷.۸.۰ (R ۲۰۰۹)*** پیاده سازی و اجرا می کنیم. نتایج عددی نشان می‌دهند که این روش برای مسائل آزمون شده بسیار موثر و کارا است.

کلمات کلیدی: بهینه سازی نامقید ؛ روش گرادیان مزدوج ؛ جستجوی خطی ؛ همگرایی سراسری.

فهرست مطالب

فصل چهارم: روش گرادیان مزدوج

فصل پنجم: روش گرادیان مزدوج با یک جستجوی خطی جدید آرمیجو-گونه

فصل ششم: نتایج عددی

فهرست الگوریتم‌ها

عنوان **صفحه**

فصل اول

بهینه سازی نامقید[1]

۱-۱ مقدمه

در این فصل ابتدا به تبیین تعاریف اولیه و مفاهیم مورد نیاز می پردازیم، سپس بهینه سازی نامقید، شرایط بهینگی و برخی از ویژگی های آن را بررسی خواهیم نمود.

۲-۱ تعاریف اولیه

تعریف ۱-۱: تابع $f:\mathbb{R}^n \to \mathbb{R}$ در نقطه ی $\bar{\mathbf{x}} \in \mathbb{R}^n$ پیوسته است، اگر برای هر $\varepsilon > 0$ وجود داشته باشد $\delta > 0$، به طوری که برای هر $\mathbf{x} \in \mathbb{R}^n$ ، $\|\mathbf{x} - \bar{\mathbf{x}}\| < \delta$ نتیجه دهد: $|f(\mathbf{x}) - f(\bar{\mathbf{x}})| < \varepsilon$.

تعریف ۱-۲: تابع پیوسته‌ی $f:\mathbb{R}^n \to \mathbb{R}$ را در نقطه‌ی $\mathbf{x}$ مشتق پذیر گویند، هرگاه تابع $g:\mathbb{R}^n \to \mathbb{R}$ موجود باشد به طوری که برای هر $0 \neq \mathbf{d} \in \mathbb{R}^n$ داشته باشیم:

[1] *Unconstrained Optimization*

$$\lim_{t\to 0}\frac{f\left(\mathbf{x}+t\mathbf{d}\right)-f\left(\mathbf{x}\right)}{t}=g^{T}\left(\mathbf{x}\right)\mathbf{d} \qquad (۱-۱)$$

بازای بردار یکه‌ی $\mathbf{d}=e_j$، اگر $\lim_{t\to 0}\frac{f\left(\mathbf{x}+te_j\right)-f\left(\mathbf{x}\right)}{t}$ موجود باشد، آنگاه آن را مشتق پاره ای، جزئی یا نسبی f نسبت به x_j گوییم و با نماد $\frac{\partial f\left(\mathbf{x}\right)}{\partial x_j}$ نشان می دهیم.

تعریف ۱-۳: تابع پیوسته‌ی $f:\mathbb{R}^n\to\mathbb{R}$ ، در نقطه‌ی $\mathbf{x}\in\mathbb{R}^n$ ، به طور پیوسته مشتق پذیر[۲] نامیده می شود، هرگاه $\left(\frac{\partial f}{\partial x_j}\right)(\mathbf{x})$ ، $j=1,2,...,n$ موجود و پیوسته باشد.

تعریف ۱-۴: تابع به طور پیوسته مشتق پذیر $f:\mathbb{R}^n\to\mathbb{R}$ ، در نقطه‌ی $\mathbf{x}\in\mathbb{R}^n$ ، به طور پیوسته مشتق پذیر مرتبه دوم[۳] نامیده می شود، هرگاه $\frac{\partial^2 f}{\partial x_i\,\partial x_j}$ ، $i,j=1,2,...,n$ موجود و پیوسته باشد.

تعریف ۱-۵: گرادیان تابع f در نقطه‌ی $\mathbf{x}$ به صورت زیر تعریف می شود:

$$\nabla f\left(\mathbf{x}\right)=\left[\left(\frac{\partial f}{\partial x_1}\right)(\mathbf{x}),\left(\frac{\partial f}{\partial x_2}\right)(\mathbf{x}),...,\left(\frac{\partial f}{\partial x_n}\right)(\mathbf{x})\right]^T \qquad (۲-۱)$$

تعریف ۱-۶: ماتریس هسیان[۴] تابع f ، به صورت یک ماتریس متقارن $n\times n$ ، با درایه های زیر تعریف می-شود:

$$\left[\nabla^2 f\left(\mathbf{x}\right)\right]_{ij}=\frac{\partial^2 f}{\partial x_i\,\partial x_j}\quad,\quad 1\le i,j\le n. \qquad (۳-۱)$$

---
[۲] *Continuously differentiable*
[۳] *Twice continuously differentiable*
[۴] *Hessian*

تعریف ۱-۷: زیر مجموعه‌ی S از فضای $\mathbb{R}^n$ را محدب گویند، هرگاه بازای هر $\alpha\in[0,1]$ و $x_1,x_2\in S$ داشته باشیم:

$$\alpha x_1+(1-\alpha)x_2\in S \tag{۱-۴}$$

تعریف ۱-۸: فرض کنید $S\subset\mathbb{R}^n$ ، یک مجموعه‌ی محدب باشد و $f:S\subset\mathbb{R}^n\to\mathbb{R}$ ، اگر برای هر $x_1,x_2\in S$ و همه ی $\alpha\in(0,1)$ داشته باشیم:

$$f\left(\alpha x_1+(1-\alpha)x_2\right)\le\alpha f\left(x_1\right)+(1-\alpha)f\left(x_2\right) \tag{۱-۵}$$

آنگاه تابع f روی مجموعه‌ی S ، محدب نامیده می شود.

تعریف ۱-۹: ماتریس متقارن $A_{n\times n}$ را نیمه معین مثبت[۵] گوییم، هرگاه برای هر $\mathbf{x}\in\mathbb{R}^n$ داشته باشیم:

$$\mathbf{x}^T A\,\mathbf{x}\ge 0$$

تعریف ۱-۱۰: ماتریس متقارن $A_{n\times n}$ را معین مثبت[۶] گوییم هرگاه برای هر $0\ne\mathbf{x}\in\mathbb{R}^n$ داشته باشیم:

$$\mathbf{x}^T A\,\mathbf{x}>0$$

تعریف ۱-۱۱: روش تکراری[۷] ، روشی است که از یک نقطه‌ی اولیه مانند $\mathbf{x}_0\in\mathbb{R}^n$ شروع می شود و دنباله‌ی تکراری $\mathbf{x}_{k+1}=H\left(\mathbf{x}_k\right)$ را تولید می کند، که در آن H رابطه‌ای بر حسب $\mathbf{x}_k$ و تکرارهای قبلی می باشد.

تعریف ۱-۱۲: گوییم یک روش تکراری دارای همگرایی سراسری[۸] است، هرگاه بازای هر نقطه‌ی شروع دلخواه $\mathbf{x}_0\in\mathbb{R}^n$، به یک جواب مسئله همگرا باشد.

---

[۵] *Positive semidefinit*

[۶] *Positive definit*

[۷] *Iterative method*

۱-۳ بهینه سازی نامقید

در این بخش، به مسئله‌ی مینیمم سازی نامقید زیر می‌پردازیم:

$$\min f\left(\mathbf{x}\right) \quad , \quad \mathbf{x}\in \mathbb{R}^n \tag{۱-۶}$$

و شرایط بهینگی را مورد بررسی قرار می‌دهیم. عموماً مینیمم کننده‌ها به دو بخش مینیمم موضعی و مینیمم سراسری تقسیم می شوند [۱].

تعریف ۱-۱۳: نقطه ی $\mathbf{x}^*$ مینمم کننده‌ی موضعی نامیده می‌شود، هرگاه $\delta>0$ وجود داشته باشد، به طوری که برای تمامی $\mathbf{x}\in \mathbb{R}^n$ ، $\left\|\mathbf{x}-\mathbf{x}^*\right\|<\delta$ داشته باشیم: $f\left(\mathbf{x}^*\right)\le f\left(\mathbf{x}\right)$.

تعریف ۱-۱۴: نقطه‌ی $\mathbf{x}^*$ مینمم کننده‌ی سراسری نامیده می شود، هرگاه برای تمامی $\mathbf{x}\in \mathbb{R}^n$ داشته باشیم: $f\left(\mathbf{x}^*\right)\le f\left(\mathbf{x}\right)$.

تعریف ۱-۱۵: فرض کنید $f:\mathbb{R}^n\to\mathbb{R}$ در $\mathbf{x}\in \mathbb{R}^n$ مشتق پذیر باشد. اگر بردار $\mathbf{d}\in \mathbb{R}^n$ موجود باشد، به طوری که $\left\langle \nabla f\left(\mathbf{x}\right),\mathbf{d}\right\rangle<0$ - ضرب نقطه‌ای بردار $\nabla f\left(\mathbf{x}\right)$ و $\mathbf{d}$ - آنگاه $\mathbf{d}$ یک مسیر کاهشی[۹] برای f در $\mathbf{x}$ نامیده می شود.

با استفاده از بسط تیلور داریم: $f\left(\mathbf{x}_k+t\mathbf{d}_k\right)=f\left(\mathbf{x}_k\right)+t\nabla^T f\left(\mathbf{x}_k\right)\mathbf{d}_k+o\left(t\right)$ ،آنگاه واضح است که اسکالر $\delta>0$ وجود دارد به طوری که بازای هر $t\in\left(0,\delta\right)$ داریم:

$$f\left(\mathbf{x}_k+t\mathbf{d}_k\right)<f\left(\mathbf{x}_k\right)$$

اگر و فقط اگر $\mathbf{d}_k$ یک مسیر کاهشی f در نقطه‌ی $\mathbf{x}_k$ باشد [۲].

[۸] *Global convergence*
[۹] *Descent direction*

تعریف ۱-۱۶: فرض کنید $\mathbf{x}\in D$، هرگاه اسکالر $\bar{\alpha}>0$ ای موجود باشد، به طوری که برای تمامی αها که $0\leq\alpha<\bar{\alpha}$، داشته باشیم $\mathbf{x}+\alpha\mathbf{d}\in D$، آنگاه بردار $\mathbf{d}$ یک مسیر شدنی[۱۰] در $\mathbf{x}$ نامیده می شود.

قضیه ۱-۱: (شرایط لازم مرتبه اول)

فرض کنید تابع $f\ : D\subset\mathbb{R}^n\to\mathbb{R}$ روی مجموعه‌ی باز D، به طور پیوسته مشتق پذیر باشد. اگر $\mathbf{x}^*\in D$ یک مینمم کننده‌ی موضعی برای (۱-۶) باشد، آنگاه $\nabla f\left(\mathbf{x}^*\right)=0$.

قضیه‌ی فوق بیان می کند که اگر $\mathbf{x}^*$ یک مینمم کننده‌ی موضعی باشد، f در $\mathbf{x}^*$ در راستای تمام جهت‌های شدنی دارای شیب صفر است و قضیه‌ی تالی نشان می‌دهد که اگر $\mathbf{x}^*$ یک مینمم کننده‌ی موضعی باشد، f در $\mathbf{x}^*$ انحنای نامنفی نزدیک به شیب صفر دارد.

قضیه ۱-۲: (شرایط لازم مرتبه دوم)

فرض کنید تابع $f\ : D\subset\mathbb{R}^n\to\mathbb{R}$ روی مجموعه‌ی باز D، به طور پیوسته مشتق پذیر مرتبه دوم باشد. اگر $\mathbf{x}^*$ یک مینمم کننده‌ی موضعی برای (۱-۶) باشد، آنگاه $\nabla f\left(\mathbf{x}^*\right)=0$ و $\nabla^2 f\left(\mathbf{x}^*\right)$ نیمه معین مثبت است.

قضیه ۱-۳: (شرایط کافی مرتبه دوم)

فرض کنید تابع $f\ : D\subset\mathbb{R}^n\to\mathbb{R}$ روی مجموعه‌ی باز D، به طور پیوسته مشتق پذیر مرتبه دوم باشد. اگر $\nabla f\left(\mathbf{x}^*\right)=0$ و $\nabla^2 f\left(\mathbf{x}^*\right)$ نیمه معین مثبت باشد، آنگاه $\mathbf{x}^*\in D$ یک مینمم کننده‌ی موضعی اکید است [۱].

---

۱۰ *Feasible direction*

۴-۱ ساختار روش‌های بهینه سازی

معمولاً روش های بهینه سازی، روش‌هایی تکراری برای یافتن مقدار مینیمم یک مسئله‌ی بهینه سازی می‌باشد. در ادامه، طرح بنیادی این گونه روش‌ها تبیین شده، نرخ همگرایی آنها مورد بررسی قرار می‌گیرد[۳].

الگوریتم ۱-۱: (الگوریتم پایه‌ای بهینه سازی)

مرحله‌ی ۰: (گام اولیه) تعیین نقطه‌ی شروع $\mathbf{x}_0 \in \mathbb{R}^n$ و تلُرانس $\varepsilon > 0$.

مرحله‌ی ۱: (شرط خاتمه) اگر $\left\|\nabla f\left(\mathbf{x}_k\right)\right\| \le \varepsilon$ توقف کن.

مرحله‌ی ۲: (یافتن مسیر) $\mathbf{d}_k$ ای پیدا کن که مسیر کاهشی باشد.

مرحله‌ی ۳: (جستجوی خطی) تعیین طول گام α_k به طوری که مقدار تابع هدف کاهش یابد، یعنی :

$$f\left(\mathbf{x}_k + \alpha_k \mathbf{d}_k\right) < f\left(\mathbf{x}_k\right).$$

مرحله‌ی ۴: (حلقه) قرار بده : $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$ و $k := k+1$ و به مرحله‌ی ۱ برو. □

شایان ذکر است، الگوریتم های مختلفی برای روش‌های بهینه سازی وجود دارند که در تعیین مسیر کاهشی و انجام جستجوی خطی با یکدیگر متفاوتند.

نرخ همگرایی یک الگوریتم می‌تواند معیاری برای سنجش تاثیر یک روش بهینه سازی باشد. فرض کنید دنباله‌ی تکرار $\left\{\mathbf{x}_k\right\}$ ، تولید شده توسط یک الگوریتم، به $\mathbf{x}^*$ همگرا شود :

$$\lim_{k \to \infty} \left\|\mathbf{x}_k - \mathbf{x}^*\right\| = 0$$

اگر عدد حقیقی $\alpha > 0$ و یک ثابت مثبت β، که مستقل از عدد تکرار k باشد، وجود داشته باشد به طوری که :

$$\lim_{k \to \infty} \frac{\left\|\mathbf{x}_{k+1} - \mathbf{x}^*\right\|}{\left\|\mathbf{x}_k - \mathbf{x}^*\right\|^{\alpha}} = \beta$$

آنگاه گوییم، $\{\mathbf{x}_k\}$ نرخ همگرایی Q از مرتبه‌ی α دارد. (نرخ همگرایی Q، به معنای نرخ همگرایی خارج قسمت[۱۱] است.)

به طور خلاصه، هنگامی که $\alpha=1$ و $\beta\in(0,1)$، گوییم دنباله‌ی $\{\mathbf{x}_k\}$ همگرایی Q –خطی دارد.
اگر $\alpha=1$ و $\beta=0$ ، یا $1<\alpha<2$ و $\beta>0$ ، گوییم دنباله‌ی $\{\mathbf{x}_k\}$ همگرایی Q –زبر خطی دارد و هنگامی که $\alpha=2$ ، گوییم $\{\mathbf{x}_k\}$ نرخ همگرایی Q –مرتبه ی دوم[۱۲] دارد.

هدف از معرفی نرخ همگرایی Q ، مقایسه‌ی سرعت همگرایی روش های مختلف است. نرخ همگرایی وابسته به α و به میزان اندکی وابسته به β است. فرض کنید دو دنباله‌ی $\{\mathbf{x}_k\}$ و $\{\mathbf{x}'_k\}$ را داشته باشیم و مرتبه‌ی Q – همگرایی آنها به ترتیب α و α' باشد، اگر $\alpha>\alpha'$، آنگاه همگرایی $Q-\alpha$ سریعتر از همگرایی دنباله‌ای با همگرایی $Q-\alpha'$ است.

معیار دیگری برای نرخ همگرایی که از نرخ همگرایی Q ضعیف‌تر است، نرخ همگرایی R می‌باشد. (نرخ همگرایی R، به معنای نرخ همگرایی ریشه[۱۳] است.)

فرض کنید $\{\mathbf{x}_k\}\subset\mathbb{R}^n$ به $\mathbf{x}^*$ همگرا شود و

$$R_p=\begin{cases}\lim\limits_{k\to\infty} Sup\left\|\mathbf{x}_k-\mathbf{x}^*\right\|^{\frac{1}{k}}, & p=1\\ \lim\limits_{k\to\infty} Sup\left\|\mathbf{x}_k-\mathbf{x}^*\right\|^{\frac{1}{p^k}}, & p>1\end{cases}$$

اگر $R_1=0$ ، گوییم $\{\mathbf{x}_k\}$ به $\mathbf{x}^*$ به صورت R –زبر خطی همگراست.

اگر $0<R_1<1$ ، گوییم $\{\mathbf{x}_k\}$ به $\mathbf{x}^*$ به صورت R –خطی همگراست. اگر $R_1=1$ ، گوییم همگرایی به صورت R –زیر خطی می باشد [۱].

---

[۱۱] *Quotient Converagence*
[۱۲] *Quadratic*
[۱۳] *Root Converagence*

فصل دوم

روش جهت‌های مزدوج

۲-۱ مقدمه

روش جهت‌های مزدوج در واقع مخلوق میانی روش کمترین شیب[۱] و روش نیوتن[۲] می‌باشد، که همگرایی را سرعت می‌بخشد و برخلاف سایر روش‌ها ی مشابه، از ارزیابی، انباره سازی و محاسبه معکوس ماتریس هسیان (یا حداقل سیستم ارزیابی مشابه) اجتناب می کند.

در این فصل روش جهت‌های مزدوج برای حل مسئله‌ی مرتبه‌ی دوم:

$$\underset{\mathbf{x}}{\text{minimize}} \left(\frac{1}{2}\mathbf{x}^T Q\ \mathbf{x} - \mathbf{b}^T \mathbf{x} \right) \qquad (۲-۱)$$

که Q ماتریس معین مثبت متقارن $n \times n$ است، مطرح و آنالیز می شود. از آنجایی که در هر مسئله‌ی غیرخطی کلی، رفتار تابع هدف در نزدیکی جواب بهینه، تقریبا مشابه یک تابع مرتبه‌ی دوم است، لذا عملکرد همگرایی مجانبی این روش روی مسائل غیرخطی شبیه رفتار همگرایی آن روی مسئله‌ی مرتبه‌ی دوم می باشد. در عمل ثابت شده است روش جهت‌های مزدوج، و به خصوص روش گرادیان مزدوج، با تابع هدف عمومی در زمره‌ی بهترین روش‌های عمومی همه منظوره مطرح است[۲].

---

۱ *Steepest descent method*
۲ *Newton method*

۲-۲ مسیرهای مزدوج

تعریف ۲-۱: ماتریس متقارن Q را در نظر بگیرید، دو بردار $\mathbf{d}_1$ و $\mathbf{d}_2$ را Q - متعامد[۳] و یا نسبت به Q مزدوج می‌نامند، اگر $\mathbf{d}_1^T Q\ \mathbf{d}_2 = 0$.

مجموعه‌ی متناهی از بردارهای $\mathbf{d}_0, \mathbf{d}_1, ..., \mathbf{d}_k$ مجموعه ی Q - متعامد می‌نامند، اگر داشته باشیم:

$$\mathbf{d}_i^T Q\ \mathbf{d}_j = 0 \qquad i, j = 0,1,\cdots,k\ ;\ i \neq j \tag{۲-۲}$$

نکته ۲-۱: اگر $Q = I$، تعریف فوق همان تعامد بردارها بر یکدیگر است.

گزاره ۲-۱: اگر Q، یک ماتریس معین مثبت باشد و مجموعه بردارهای ناصفر $\mathbf{d}_0, \mathbf{d}_1, ..., \mathbf{d}_k$، Q - متعامد باشند، آنگاه این بردارها مستقل خطی اند[۲].

برهان. فرض کنید ثابت های α_j بازای $j = 0,1,...,k$ موجود باشند به طوری که:

$$\alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \cdots + \alpha_k \mathbf{d}_k = 0$$

برای i دلخواه، با ضرب $\mathbf{d}_i^T Q$ از سمت چپ در رابطه ی فوق داریم:

$$\alpha_i \mathbf{d}_i^T Q\ \mathbf{d}_i = 0$$

به دلیل معین مثبت بودن Q داریم:

$$\mathbf{d}_i^T Q\ \mathbf{d}_i > 0$$

پس $\alpha_i = 0$. □

---

[۳] *Q-orthogonal*

اکنون به بررسی اهمیت استفاده از مفهوم Q - متعامد برای حل مسئله‌ی مرتبه‌ی دوم (۲-۱) می پردازیم. چون Q معین مثبت است، می‌دانیم جواب بهینه‌ی مسئله‌ی (۲-۱)، یکتا بوده و همان جواب یکتای دستگاه:

$$Q\mathbf{x} = \mathbf{b} \tag{۲-۳}$$

می باشد. از این رو مسئله‌ی مینیمم سازی مرتبه‌ی دوم، معادل حل دستگاه معادلات خطی است.

فرض کنید ماتریس $Q_{n\times n}$، معین مثبت و بردارهای ناصفر $\mathbf{d}_{n-1},\ldots,\mathbf{d}_1,\mathbf{d}_0$ ، Q - متعامد باشند. استقلال خطی این بردارها موجب می شود که $\alpha_{n-1},\cdots,\alpha_0$ موجود باشند، به طوری که جواب $\mathbf{x}^*$ برای (۲-۱) یا (۲-۳) به صورت زیر گسترانیده شود:

$$\mathbf{x}^* = \alpha_0\mathbf{d}_0 + \alpha_1\mathbf{d}_1 + \cdots + \alpha_{n-1}\mathbf{d}_{n-1} \tag{۲-۴}$$

برای i دلخواه، با ضرب $\mathbf{d}_i^T Q$ از سمت چپ در رابطه‌ی فوق مستقیماً نتیجه می‌شود:

$$\alpha_i = \frac{\mathbf{d}_i^{\,T} Q\ \mathbf{x}^*}{\mathbf{d}_i^{\,T} Q\ \mathbf{d}_i} = \frac{\mathbf{d}_i^{\,T}\mathbf{b}}{\mathbf{d}_i^{\,T} Q\ \mathbf{d}_i} \tag{۲-۵}$$

رابطه‌ی فوق نشان می دهد که α_i ها و به تبع آن $\mathbf{x}^*$ با محاسبه‌ی ضرب های اسکالر ساده‌ای بدست می‌آیند؛ بنابراین نتیجه‌ی نهایی به صورت زیر است:

$$\mathbf{x}^* = \sum_{i=0}^{n-1} \frac{\mathbf{d}_i^{\,T}\mathbf{b}}{\mathbf{d}_i^{\,T} Q\ \mathbf{d}_i}\mathbf{d}_i \tag{۲-۶}$$

در رابطه‌ی (۲-۶) دو ایده‌ی اساسی نهفته است، اول انتخاب یک مجموعه‌ی متعامد از $\mathbf{d}_i$ ها به قسمی که با در نظر گرفتن یک حاصلضرب اسکالری مناسب، همه جملات سمت راست رابطه‌ی (۲-۴) به جز جمله‌ی i-اُم صفر شوند.

مزیت دیگر استفاده از بردارهای Q - متعامد برای بدست آوردن رابطه‌ای برای α_i ها است، که آن‌ها را بر حسب بردار معلوم $\mathbf{b}$ به جای بردار نامعلوم $\mathbf{x}^*$ معین می کند. یعنی بدون در دست داشتن $\mathbf{x}^*$ می‌توان ضرایب را محاسبه کرد و با این روند، بسط $\mathbf{x}^*$ را در فرایندی n مرحله‌ای، که در مرحله‌ی i-اُم، $\alpha_i\mathbf{d}_i$ اضافه می‌شود، به دست آورد.

قضیه ۲-۱: (قضیه ی مسیرهای مزدوج)

فرض کنید $\{\mathbf{d}_i\}_{i=0}^{n-1}$، مجموعه‌ای از بردارهای ناصفر Q - متعامد باشد. برای هر $\mathbf{x}_0 \in \mathbb{R}^n$ دنباله‌ی $\{\mathbf{x}_n\}$ تولید شده از:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \quad , \quad k \geq 0 \tag{۲-۷}$$

با:

$$\alpha_k = -\frac{g_k^T \mathbf{d}_k}{\mathbf{d}_k^T Q \ \mathbf{d}_k} \tag{۲-۸}$$

و

$$g_k = Q\mathbf{x}_k - \mathbf{b} \tag{۲-۹}$$

پس از حداکثر n تکرار، به جواب یکتای $\mathbf{x}^*$ از $Q\mathbf{x} = \mathbf{b}$ همگرا می‌شود، یعنی $\mathbf{x}_n = \mathbf{x}^*$ [۲].

برهان. فرض کنید $\mathbf{x}^*$ جواب یکتای $Q\mathbf{x} = \mathbf{b}$ باشد، از آنجایی که $\mathbf{d}_k$ها مستقل خطی اند، $\alpha_{n-1}, \cdots, \alpha_0$ موجودند، به طوری که:

$$\mathbf{x}^* - \mathbf{x}_0 = \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \cdots + \alpha_{n-1} \mathbf{d}_{n-1} \tag{۲-۱۰}$$

همان طور که برای به دست آوردن رابطه‌ی (۲-۵) گفتیم، با ضرب در $\mathbf{d}_i^T Q$ از سمت چپ به رابطه‌ی زیر می‌رسیم:

$$\alpha_k = \frac{\mathbf{d}_k^T Q \left(\mathbf{x}^* - \mathbf{x}_0\right)}{\mathbf{d}_k^T Q \ \mathbf{d}_k} \tag{۲-۱۱}$$

حال با استفاده از فرایند تکرار (۲-۷) از $\mathbf{x}_0$ تا $\mathbf{x}_k$ داریم:

$$\mathbf{x}_k - \mathbf{x}_0 = \alpha_0 \mathbf{d}_0 + \alpha_1 \mathbf{d}_1 + \cdots + \alpha_{k-1} \mathbf{d}_{k-1} \tag{۲-۱۲}$$

و از Q - متعامد بودن بردارهای $\mathbf{d}_k$ می‌توانیم بنویسیم:

$$\mathbf{d}_k^T Q \left(\mathbf{x}_k - \mathbf{x}_0\right) = 0 \tag{۲-۱۳}$$

با اضافه نمودن $\mathbf{d}_k{}^T Q\left(\mathbf{x}_k - \mathbf{x}_0\right)$ به صورت کسر در رابطه‌ی (۲-۱۱) داریم:

$$\alpha_k = \frac{\mathbf{d}_k{}^T Q\left(\mathbf{x}^* - \mathbf{x}_k\right)}{\mathbf{d}_k{}^T Q\ \mathbf{d}_k} = \frac{\mathbf{d}_k{}^T\left(Q\mathbf{x}^* - Q\mathbf{x}_k\right)}{\mathbf{d}_k{}^T Q\ \mathbf{d}_k}$$
$$= \frac{\mathbf{d}_k{}^T\left(\mathbf{b} - Q\mathbf{x}_k\right)}{\mathbf{d}_k{}^T Q\ \mathbf{d}_k} = -\frac{g_k^T \mathbf{d}_k}{\mathbf{d}_k{}^T Q\ \mathbf{d}_k}$$

که همسان با (۲-۸) می‌باشد. □

۲-۳ خواص کاهشی روش جهت‌های مزدوج

β_k را بعنوان زیر فضای $\mathbb{R}^n$، تولید شده توسط $\{\mathbf{d}_0, \mathbf{d}_1, \ldots, \mathbf{d}_{k-1}\}$ در نظر می‌گیریم. می‌بایست نشان دهیم که در روند اجرای روش جهت‌های مزدوج، هر $\mathbf{x}_k$ تابع هدف را بر روی چند وجهی k بعدی $\mathbf{x}_0 + \beta_k$ مینیمم می‌کند.

لم ۲-۱: فرض کنید f به طور پیوسته مشتق پذیر باشد، f روی مجموعه‌ی محدب D، محدب است اگر و فقط اگر:

$$f\left(\mathbf{y}\right) \geq f\left(\mathbf{x}\right) + \nabla f^T\left(\mathbf{x}\right)\left(\mathbf{y} - \mathbf{x}\right) \qquad (۲\text{-}۱۴)$$

لم ۲-۲: فرض کنید f به طور پیوسته مشتق پذیر و تابعی محدب روی مجموعه‌ی محدب D باشد. اگر نقطه‌ی $\mathbf{x}^* \in D$ موجود باشد به طوری که برای هر $\mathbf{y} \in D$،

$$\nabla f^T\left(\mathbf{x}^*\right)\left(\mathbf{y} - \mathbf{x}^*\right) \geq 0$$

آنگاه $\mathbf{x}^*$ نقطه‌ی مینیمم کننده‌ی سراسری تابع f روی D است.

برهان. از آنجایی که $\mathbf{y}-\mathbf{x}^*$ یک جهت شدنی در $\mathbf{x}^*$ است، شرایط معین شده در حکم شرایط لازم مرتبه اول گفته شده در بخش ۱-۳ می‌باشد و اثبات بی واسطه از لم ۲-۱ به دست می‌آید:

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \nabla f^T(\mathbf{x}^*)(\mathbf{y}-\mathbf{x}^*) \geq f(\mathbf{x}^*). \quad \square$$

قضیه ۲-۲: (قضیه‌ی زیر فضای انبساطی[۴])

فرض کنید $\{\mathbf{d}_i\}_{i=0}^{n-1}$ دنباله‌ای غیر صفر از بردارهای Q - متعامد در $\mathbb{R}^n$ باشد. در این صورت بازای هر $\mathbf{x}_0 \in \mathbb{R}^n$ دنباله‌ی $\{\mathbf{x}_k\}$ تولید شده از رابطه‌ی

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \tag{۲-۱۵}$$

$$\alpha_k = -\frac{g_k^T \mathbf{d}_k}{\mathbf{d}_k^T Q\, \mathbf{d}_k} \qquad (\mathrm{g_k} = Q\mathbf{x} - \mathbf{b}) \tag{۲-۱۶}$$

دارای این خاصیت است که $\mathbf{x}_k$، مینیمم کننده‌ی تابع

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\, \mathbf{x} - \mathbf{b}^T \mathbf{x}$$

روی خط $\mathbf{x} = \mathbf{x}_{k-1} + \alpha \mathbf{d}_{k-1}$، $(-\infty < \alpha < \infty)$ و همچنین روی چند وجهی خطی $\mathbf{x}_0 + \beta_k$ می باشد[۲].

برهان. کافی است نشان دهیم $\mathbf{x}_k$ مینیمم کننده‌ی f روی چند وجهی $\mathbf{x}_0 + \beta_k$ است، زیرا در برگیرنده‌ی خط $\mathbf{x} = \mathbf{x}_{k-1} + \alpha \mathbf{d}_{k-1}$ نیز می‌باشد. از آنجایی که f تابعی اکیداً محدب است، اگر نشان دهیم که g_k عمود بر β_k می‌باشد، (به عبارت دیگر گرادیان f در نقطه‌ی $\mathbf{x}_k$، بر زیر فضای β_k عمود است.) آنگاه طبق لم ۲-۲ حکم برقرار است، وضعیت در شکل ۲-۱ نشان داده شده است.

---
۴ *Expanding Subspace Theorem*