مقایسه روش های تکراری نیوتن، بریدن و شبه‌نیوتن برای حل دستگاههای معادلات غیرخطی

هاشم صابری نجفی[1] - سهراب کرد رستمی[2] - نرگس سهرابی گیلانی[2]

[1]گروه ریاضی ، دانشکده علوم پایه ، دانشگاه گیلان
[2]گروه ریاضی ، دانشگاه آزاد اسلامی واحد لاهیجان

چکیده

در این مقاله روش های تکراری نیوتن، شبه‌نیوتن و بریدن از لحاظ ساختار- کاربرد والگوریتم مورد مطالعه قرار گرفته‌اند. این سه روش را بر روی مثالهایی از حیث همگرایی - سرعت - دقت مورد بررسی قرار دادیم. نتایج را در جدولها و نمودارهایی ارایه کرده‌ایم. همچنین روش های نیوتن و شبه نیوتن را برای روش های بهینه سازی نامقید و با استفاده از مشتق دوم مورد مطالعه قرار می‌دهیم و الگوریتمی را معرفی می‌کنیم که بدون از دست دادن مزیت‌هایی نظیر پایداری و توانایی نگهداری وضعیت، تعداد اعمال حسابی معنی‌دار را کم نموده و دیگر نیازی به جمع آوری Q در حین فرایند تجزیه در هر تکرار نخواهد بود.

کلمات کلیدی: شبه نیوتن، بهینه سازی نامقید، تجزیه LU، تجزیه PR

۱ مقدمه

درعلوم و مهندسی با مسائلی روبرو هستیم که حل آنها به دستگاه معادلات غیر خطی

$$\begin{aligned} &f: R^n \to R^n \\ &f(x) = 0 \end{aligned} \qquad (۱-۱)$$

منجر می‌شود که با توجه به ضرایب معادلات دستگاه فوق و بعد آن می‌توان از روش های تکراری برای حل چنین دستگاهی استفاده کرد. در این مقاله روش های تکراری نیوتن، شبه نیوتن و بریدن را از لحاظ ساختار مورد مطالعه قرار می‌دهیم. در ابتدا به تعریف همگرایی R- زبرخطی و همگرایی Q- زبرخطی می‌پردازیم.

۱-۱ **همگرایی R-زبرخطی و همگرایی Q-زبرخطی**- اگر دنباله $\{x_k\}$ به x^* همگرا باشد آنگاه

$\{x_k\}$, R- زبرخطی به x^* همگرا است اگر

$$\lim_{k\to\infty} \left\| x_{k+1} - x^* \right\|^{\frac{1}{k}} = 0$$

و دنباله $\{x_k\}$, Q- زبرخطی به x^* همگرا است، اگر یک دنباله $\{\alpha_k\}$ همگرا به صفر، موجود باشد به طوریکه

$$\left\|x_{k+1}-x^*\right\| \le \alpha_k \left\|x_k - x^*\right\|$$

واضح است که همگرایی Q- زبرخطی، همگرایی R- زبرخطی را ایجاب می‌کند ولی همگرایی را نگه می‌دارد.[۶]

۲ ساختار روش نیوتن

دستگاه معادلات غیرخطی

$$f : R^n \to R^n \qquad (۱-۲)$$
$$f(x)=0$$

را به روش تقریب‌های متوالی حل می‌کنیم. p امین تخمین ریشه بصورت $x^{(p)}=\left(x_1^{(p)},\dots,x_n^{(p)}\right)$ می‌باشد و ریشه دقیق دستگاه را بصورت

$$x = x^{(p)} + \varepsilon^{(p)} \qquad (۲-۲)$$

نمایش می‌دهیم که ε^{ν} خطای تقریب ریشه با $x^{(p)}$ می‌باشد. اگر f به طور پیوسته مشتق‌پذیر باشد با توجه به تقریب بسط تیلور می‌توان نوشت:

$$f\left(x^{(p)}+\varepsilon^{(p)}\right) \approx f\left(x^{(p)}\right) + f'\left(x^{(p)}\right)\varepsilon^{(p)} \qquad (۳-۲)$$

چون $f\left(x^{(p)}+\varepsilon^{(p)}\right)=0$ و با فرض نامنفرد بودن $J\left(x^{(p)}\right)$ خواهیم داشت:

$$\varepsilon^{(p)} = -J^{-1}\left(x^{(p)}\right) f\left(x^{(p)}\right)$$

و با توجه به رابطه فوق و رابطه (۲-۲) طرح تکراری زیر برای یافتن جواب ارایه می‌شود:

$$x^{(p+1)} = x^{(p)} - J^{-1}\left(x^{(p)}\right) f\left(x^{(p)}\right) \qquad (۴-۲)$$
$$p = 0,1,2,\dots$$

قضیه ۲-۱ وجود ریشه‌های یک دستگاه و همگرایی روش نیوتن- دستگاه غیرخطی (۲-۱) داده شده که در آن $f(x)$ پیوسته و $f(x)\in C^{(2)}(\omega)$ باشد و $x^{(0)}$ همراه با $\mathrm{H}-$همسایگی بسته‌اش یعنی

$$\overline{U}_{\mathrm{H}}\left(x^{(0)}\right) = \left\{ \left\|x - x^{(0)}\right\| \le \mathrm{H} \right\}$$

متعلق به ω باشند و اگر

۱- ماتریس ژاکوبی دارای معکوس $\Gamma_0 = J^{-1}\left(x^{(0)}\right)$ برای $x = x^{(0)}$ می‌باشد که $\|\Gamma_0\| \le A_0$

۲- $\left\|\Gamma_0 f\left(x^{(0)}\right)\right\| \le \mathrm{B}_0 \le \frac{\mathrm{H}}{2}$

۳- $\sum_{k=1}^{n}\left|\frac{\partial^2 f_i(x)}{\partial x_j \partial x_k}\right| \le C \qquad : \ x \in \overline{U}_{\mathrm{H}}\left(x^{(0)}\right) \qquad i,j=1,2,\ldots,n$

۴- ثابت‌های C , B_0 , A_0 در نامعادله $\mu_0 = 2n\mathrm{A}_0\mathrm{B}_0 C \le 1$ صدق کند.

آنگاه فرایند نیوتن (۲-۴) با تقریب اولیه $x^{(0)}$ همگرا خواهد شد و $x^* = \lim_{p\to\infty} x^{(p)}$ یک جواب دستگاه $f(x)=0$ است چنانکه

$$\left\|x^* - x^{(0)}\right\| \le 2\mathrm{B}_0 \le \mathrm{H}$$

و برای تخمین‌های متوالی $p=0,1,\ldots$ $x^{(p)}$ در نامعادله:

$$\left\|x^* - x^{(p)}\right\| \le \left(\frac{1}{2}\right)^{p-1} \mu_0^{2^p-1}\mathrm{B}_0$$

صادق است و برای $\mu_0 < 1$ همگرایی روش نیوتن زبرخطی می‌باشد.[1]

۳ روش نیوتن اصلاح شده

یک مشکل اساسی در شکل فرایند نیوتن آن است که در هر مرحله ملزم به محاسبه ماتریس ژاکوبی می‌باشیم. حال اگر $J^{-1}(x)$ در همسایگی جواب x^* وتخمین اولیه $x^{(0)}$ به قدر کافی نزدیک به x^*، پیوسته باشد می‌توان $J^{-1}\left(x^{(p)}\right) \approx J^{-1}\left(x^{(0)}\right)$ را تخمین زد که در این صورت به روش نیوتن اصلاح شده می‌رسیم:

$$\xi^{(p+1)} = \xi^{(p)} - J^{-1}\left(x^{(0)}\right) f\left(x^{(p)}\right) \qquad p=0,1,2,\ldots$$

که $\xi^{(0)} = x^{(0)}$. دقت کنید که تخمین اولیه $\xi^{(1)}$, $x^{(1)}$ با توجه به فرمول‌های نیوتن و نیوتن اصلاح شده عبارت است از $x^{(1)} = \xi^{(1)}$.

۴ الگوریتم روش نیوتن

$TOL \in R$, $N \in R$, $x_0 \in R^n$ داده شده است.

گام ۱. قرار دهید: $K=0$, $P=1$

گام ۲. برای $P \le N$ مراحل (۳) تا (۷) را اجرا کنید.

گام ۳. $F(x_K)$, $J(x_K)$ را محاسبه کنید.

گام ۴. معکوس ژاکوبین $J(x_K)$ را محاسبه کنید. $\left(J^{-1}(x_K)\right)$

گام ۵. قرار $x_{K+1} = x_K - J^{-1}(x_K) F(x_K)$ دهید:

گام ۶. اگر $\left\|F(x_{K+1})\right\| < TOL$ مقدار x_{K+1} را چاپ کنید و به (۸) بروید.

گام ۷. قرار دهید: $K=K+1$, $P=P+1$

گام ۸. پایان الگوریتم

۵ ساختار روش بریدن[1]

دستگاه (۲-۱) را در نظر می گیریم. ساده‌ترین فرم روش بریدن بصورت

$$x_{k+1} = x_k - B_k^{-1} f(x_k) \qquad k = 0,1,2,\ldots \tag{۱-۵}$$

می‌باشد که B_k تقریب ماتریس ژاکوبین در x_k است.[6]

همسایگی کوچکی حول x_k در نظر می گیریم:

$$f(x) \approx f(x_k) + J(x_k)(x_k - x)$$

اگر $x = x_{k+1}$ ، $y_k = f(x_{k+1}) - f(x_k)$ و B_k تقریب $J(x_k)$ باشد، خواهیم داشت:

$$y_k = B_k(x_k - x_{k+1}) \tag{۲-۵}$$
$$\Rightarrow x_{k+1} = x_k - B_k^{-1} y_k$$

فرمول فوق وضعیت شبه نیوتن نامیده می‌شود.

بریدن یک تقریب B_k برای $J(x_k)$ می‌باشد چنانکه B_{k+1} تقریب $J(x_{k+1})$ از B_k با $O(n^3)$ عمل حسابی در هر تکرار باشد. به غیر از حالت $n=1$, B_{k+1} منحصربه‌فرد تعریف نمی‌شود و بریدن با دلیل ثابت کرد که هیچ دلیل واقعی برای اختلاف داشتن B_{k+1} از B_k روی متمم متعامد S_k وجود ندارد.[3]

بنا براین درا ین روش B_{k+1} طوری انتخاب می‌شود که تغییر در f بوسیله B_{k+1} در جهت Z عمود بر S_k قابل پیش‌بینی است ونظیر پیش‌بینی B_k می‌باشد. این مطلب به

$$B_{k+1}Z = B_k Z \qquad \langle Z, S_k \rangle = 0 \tag{۳-۵}$$

منجر می‌شود. از (۵-۲) و (۵-۳) ماتریس‌های $\{B_k\}$ با

$$B_{k+1} = B_k + \frac{(y_k - B_k S_k) S_k^T}{\|S_k\|^2} \tag{۴-۵}$$

$$S_k = x_{k+1} - x_k$$
$$y_k = f(x_{k+1}) - f(x_k) \tag{۵-۵}$$

تعریف می‌شود.

درحل به روش بریدن با استفاده از فرمول‌های (۵-۴) و (۵-۱) تعداد ارزیابی اعمال اسکالر از $n^2 + n$ به n کاهش داده می‌شود اما برای کاهش پیچیدگی از $O(n^3)$ به $O(n^2)$ از فرمول شرمن - موریسن - وودباری برای محاسبه وارون $\{B_k\}$ استفاده می‌شود که خواهیم داشت:

[1] .Broyden

$$B_{k+1}^{-1} = B_k^{-1} + \left(\frac{S_k - B_k^{-1} y_k}{S_k^T B_k^{-1} y_k} \right) S_k^T B_k^{-1} \qquad (۶-۵)$$
$$S_k^T B_k^{-1} y_k \neq 0$$

این روش در بیشترحالات خوب کار می‌کند ولی دارای این مشکل است که بررسی بد وضع بودن B_k سخت است.

قضیه ۵-۱ فرض کنید $f : R^n \to R^n$ در مجموعه باز محدب D به طور پیوسته مشتق پذیر باشد و فرض کنید $f(x^*) = 0$ و برای $x^* \in D$, $f'(x^*)$ نامنفرد باشد و با فرض اینکه f در شرط لیپشتس در x^* صدق کند، روش بریدن تعریف شده توسط (۵-۴) و (۵-۵) و (۵-۲) را در نظر بگیرید. آنگاه روش بریدن به طور موضعی و Q- زبرخطی به x^* همگرا است.[7]

دقت شود که همگرایی کلی برای توابع غیرخطی صادق نیست. در حالت یک بعدی روش بریدن ضرورتا به روش قطعه‌ای کاهش می‌یابد و این روش می‌تواند دوری باشد.

۶ الگوریتم روش بریدن

$x_0 \in R^n \quad , \quad N \in R \quad , \quad \varepsilon \in R$ داده شده است.

گام ۱. $f(x_0)$, $J(x_0)$ را محاسبه کنید.

گام ۲. $p = 1 \quad , \quad k = 0$

گام ۳. $B_0 = J^{-1}(x_0)$

گام ۴. برای $(p \le N)$ مراحل (۵) تا (۱۰) را اجرا کنید.

گام ۵. قرار دهید $S_k = -B_k^{-1} f(x_k)$ و S_k را محاسبه کنید.

گام ۶. $x_{k+1} = x_k + S_k$.

گام ۷. اگر $\varepsilon > \| f(x_{k+1}) \|$ مقدار x_{k+1} را چاپ کنید و به (۱۱) بروید.

گام ۸. قرار دهید: $y_k = f(x_{k+1}) - f(x_k)$

گام ۹. قرار دهید: $B_{k+1}^{-1} = B_k^{-1} + \left(\frac{S_k - B_k^{-1} y_k}{S_k^T B_k^{-1} y_k} \right) S_k^T B_k^{-1}$

گام ۱۰. به k یک وا حد اضافه کنید.

گام ۱۱. پایان الگوریتم

۷ ساختار روش شبه نیوتن

فرض کنیم دستگاه (۱-۱) را داشته باشیم. فرم روش شبه نیوتن بصورت زیر است:

$$x_{k+1} = x_k + H_k f(x_k) \qquad k = 0,1,2,\ldots \qquad (۱-۷)$$

که H_k تقریب وارون ماتریس ژاکوبین است و در این روش H_k منحصر بفرد تعریف می‌شود و همواره در $H_k = -B_k^{-1}$ صدق می‌کند که B_k تقریب ماتریس ژاکوبین می‌باشد.

روش شبه نیوتن به دلیل برقراری شرایط

$$H_{k+1}V_k = H_kV_k \qquad \langle V_K, Z\rangle = 0 \tag{۲-۷}$$

به دست آمده است. ماتریسهای $\{H_k\}$ با استفاده از

$$H_{k+1} = H_k - \frac{(S_k + H_k y_k)y_k^T}{y_k^T y_k}$$

که

$$S_k = x_{k+1} - x_k \qquad y_k = f(x_{k+1}) - f(x_k)$$

تعریف می‌شوند و با استفاده از فرمول شرمن- موریسن- وودباری می‌توان وارون $\{H_{k+1}\}$ را به طورمستقیم ازوارون $\{H_k\}$ به دست آورد که خواهیم داشت:

$$x_{k+1} = x_k - B_k^{-1} f(x_k) \qquad k = 0,1,2,\ldots \tag{۳-۷}$$

به طوریکه

$$B_{k+1} = B_k + \left(\frac{(y_k - B_k S_k)y_k^T}{y_k^T (B_k) S_k}\right)(B_k) \qquad S_k = x_{k+1} - x_k \;,\; y_k = f(x_{k+1}) - f(x_k)$$

۸ الگوریتم روش شبه نیوتن

$x_0 \in R^n \;,\; N \in R \;,\; \varepsilon \in R$ داده شده است.

گام ۱. قرار دهید: $p = 1 \;,\; k = 0$.

گام ۲. قرار دهید: $J(x_0) = B_0$ و $f(x_0)$ را محاسبه کنید.

گام ۳. برای $(p \le N)$ مراحل (۴) تا (۹) را اجرا کنید.

گام ۴. قرار دهید $B_k S_k = -f(x_k)$ و S_k را محاسبه کنید.

گام ۵. $x_{k+1} = x_k + S_k$.

گام ۶. اگر $\|f(x_{k+1})\| < \varepsilon$ مقدار x_{k+1} را چاپ کنید و به (۱۰) بروید.

گام ۷. قرار دهید: $y_k = f(x_{k+1}) - f(x_k)$.

گام ۸. قرار دهید: $B_{k+1}^{-1} = B_k^{-1} + \left(\frac{S_k - B_k^{-1} y_k}{S_k^T B_k^{-1} y_k}\right) S_k^T B_k^{-1}$.

گام ۹. به $p \;,\; k$ یک واحد اضافه کنید.

گام ۱۰. پایان الگوریتم.

۹ مثال عددی

در این بخش با چند مثال عددی الگوریتم‌ها را بررسی می‌کنیم و در جداول و نمودارها این سه الگوریتم را از نقطه‌نظر همگرایی، سرعت و دقت مورد بررسی قرار می‌دهیم. منظور از خطا $\|f(x_{k+1})\| < TOL$ می‌باشد.

تذکر: در نمودارها محور x ها نمایشگر گام‌ها و محور y ها نمایانگر خطای گام مربوطه می‌باشد.

مثال ۱ دستگاه معادلات غیرخطی

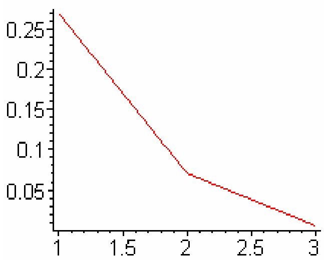
$$\begin{cases} 2x-3y+z-4=0 \\ 2x+y-z+4=0 \\ x^2+y^2+z^2-4=0 \end{cases}$$

را با نقطه اولیه $(-0.5,-1.5,1.5)^t$ در نظر بگیرید. سه الگوریتم را برای مثال فوق بکارگرفتیم و نتایج در جدول شماره ۱ و نمودارهای (۱–۱) و (۱–۲) و (۱–۳) قابل مشاهده می‌باشد.

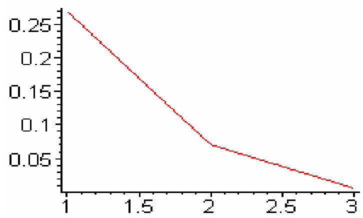
Iteration error	شبه نیوتن	برید ن	نیوتن	
E(1)	0.07250000000	0.07250000000	0.07250000000	
E(2)	0.001565798000	0.004996888000	0.004996888000	
E(3)	0.00009857700002	0.00003193100003	0.00003193100003	$\|AX-b\|$
E(4)	$0.6479000000\ 10^{-5}$	$0.2000000000\ 10^{-8}$	$0.2000000000\ 10^{-8}$	
E(5)	$0.4250000000\ 10^{-6}$			
E(6)	$0.2903446228\ 10^{-7}$			
E(7)	$0.1000000000\ 10^{-8}$			

جدول ۱: خطا برای دستگاه اول

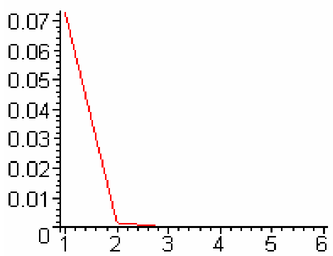
در جدول۱، $E(i)$ که $i=1,2,\dots,7$ ، خطای مرحله i–ام است.



۱–۱: نمودار خطا (روش نیوتن)



۱–۲: نمودار خطا (روش شبه نیوتن)



۱–۳: نمودار خطا (روش بریدن)

مثال ۲ دستگاه معادلات غیرخطی زیر با نقطه اولیه $(0.5,1)^t$ داده شده است.

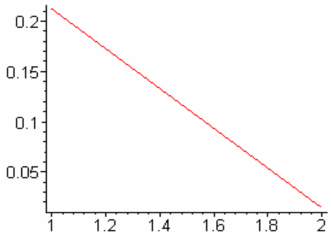
$$\begin{cases} x^2+y^2-4x=0 \\ y^2+2x-2=0 \end{cases}$$

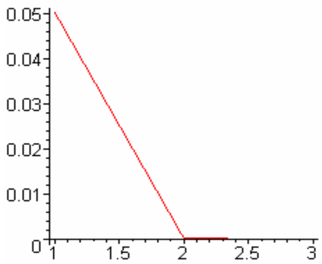
نتایج مقایسه سه الگوریتم در جدول شماره ۲ و نمودارهای (۲-۱) و (۲-۲) و (۲-۳) قابل مشاهده می‌باشد.

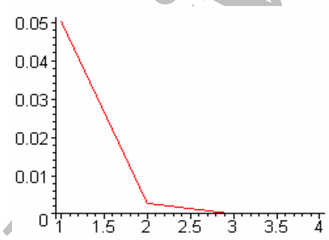
Iteration error	شبه نیوتن	برید ن	نیوتن	
E(1)	0.05031152949	0.05031152949	0.05031152949	
E(2)	0.002794625970	0.0002698000585	0.0002698000585	
E(3)	0.00001545728650	0.9899494937 10^{-8}	0.9899494937 10^{-8}	$\lVert AX-b \rVert$
E(4)	0.8765842800 10^{-7}			
E(5)	0.			

جدول ۲ : خطا برای دستگاه دوم

در جدول ۲، $E(i)$ که $i=1,2,\ldots,5$ ، خطای مرحله $i-$ام است.







۲-۱: نمودار خطا (روش نیوتن)

۲-۲: نمودار خطا (روش شبه نیوتن)

۲-۳: نمودار خطا (روش بریدن)

مثال ۳ دستگاه معادلات غیرخطی

$$\begin{cases} x^2-2x-y+1=0 \\ x^2+y^2-1=0 \end{cases}$$

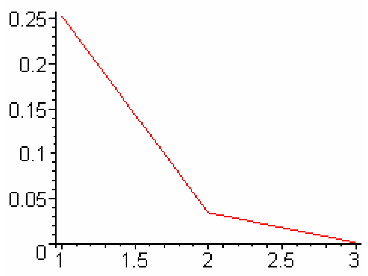
را با نقطه اولیه $(0.9,0.2)^t$ در نظر بگیرید. سه الگوریتم را برای مثال فوق بکار گرفتیم، نتایج درجدول شماره ۳ و نمودارهای (۳-۱) و (۳-۲) و (۳-۳) قابل مشاهده می‌باشد.

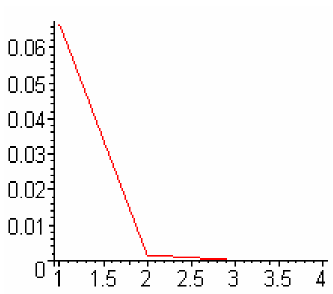
$\|AX-b\|$

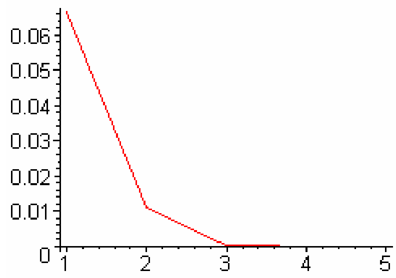
Iteration error	شبه نیوتن	برید ن	نیوتن
E(1)	0.06632773745	0.06632773745	0.06632773745
E(2)	0.01085931580	0.001517247582	0.001517247582
E(3)	0.0001416305030	$0.1295772341\ 10^{-5}$	$0.1295772341\ 10^{-5}$
E(4)	$0.3297180044\ 10^{-5}$	$0.1000000000\ 10^{-9}$	$0.1000000000\ 10^{-9}$
E(5)	$0.9588013350\ 10^{-7}$		
E(6)	$0.1000000000\ 10^{-9}$		

جدول۳: خطا برای دستگاه سوم

در جدول۳، $E(i)$ که $i=1,2,\ldots,6$ ، خطای مرحله $i-$ ام است.







۳-۳. نمودار خطا (روش بریدن) **۳-۲.** نمودار خطا (روش شبه نیوتن) **۳-۱.** نمودار خطا (روش نیوتن)

۱۰ توسعه روش های نیوتن و شبه نیوتن با تعیین طول گام و جهت جدید

مساله بهینه‌سازی نامقید بفرم $\begin{matrix} \varphi : R^n \to R \\ \min\limits_{x} \varphi(x) \\ x \in R^n \end{matrix}$ می‌باشد. اگر علاوه بر بردارگرادیان از مشتق دوم تابع هدف $\varphi(x)$ نیز استفاده شود همگرایی سریعترخواهد شد. روش نیوتن تکرار جدید x_{k+1} را با مینیمم کردن

$$q_k(s_k)=\varphi(x_k+s_k)\approx\varphi(x_k)+g(x_k)^T s_k+\frac{1}{2}s_k^T H(x_k)s_k$$

وبا استفاده از تابع $\varphi(x)$ در تکرار جاری x_k تعیین می‌کند. و زمانی که نقطه اولیه x_0 تخمین مناسبی از ریشه نباشد روش نیاز به اصلاح دارد.

توجه شود که اگر گام نیوتن $g_k^T H(x_k)^{-1} g_k \le 0$ باشد، جهت کاهشی نخواهد بود و این بخاطر خاصیت معین مثبت بودن یا حداقل معین منفی نبودن $H(x^*)$ در مجاورت نقطه x^* (مینیمم موضعی) رخ نمی‌دهد. این دلیل انتخاب بردار گرادیان به عنوان جهت می‌باشد زیرا احتمال آنکه گام نیوتن به نقطه‌ای زینی منجر شود وجود دارد. بسط سری تیلور تابع گرادیان، حول نقطه x_k و طول s_k، بصورت

$$g(x_k + s_k) = g_k + H(x_k)s_k + \cdots \tag{۱-۱۰}$$

می‌باشد. در روش شبه‌نیوتن B_k را بعنوان تخمین ماتریس هسین در $k-$امین مرحله بکار می‌گیریم که B_{k+1} خمیدگی φ را در طول $s_k = x_{k+1} - x_k$ تخمین می‌زند یعنی

$$B_{k+1}s_k = \gamma_k$$
$$\gamma_k = g(x_{k+1}) - g(x_k)$$

جائیکه γ_k در گرادیان تغییر می‌کند که شرط شبه‌نیوتن نامیده می‌شود.
شرط شبه‌نیوتن می‌تواند با استفاده از

$$B_{k+1} = B_k + \frac{r_k s_k^T + s_k r_k^T}{s_k^T s_k} - \frac{(r_k^T s_k)s_k s_k^T}{(s_k^T s_k)^2}$$

یعنی (PSB)[۲] به روز شود که $r_k = \gamma_k - B_k s_k$. ماتریس به روز شده $B_{k+1} - B_k$ اصلاح شده مرتبه ۲ است. نیز معادله دیگری به نام (BFGS)[۳]، برای به روز کردن B_k یافته شده است که عبارت است از

$$B_{k+1} = B_k - \frac{B_k Y_k Y_k^T B_k}{Y_k^T B_k Y_k} + \frac{\gamma_k \gamma_k^T}{\gamma_k^T Y_k}$$

و می‌تواند گزینه بهتری از مورد (PSB) باشد.

پیچیدگی روش $O(n^3)$ خواهد بود. هرچند چون B_k بهبود یافته مرتبه دوم B_{k-1} می‌باشد امکان حل دستگاه به روشی کاراتر وجود دارد. یک روش محاسبه تخمین B_k^{-1} استفاده از فرمول شرمن-موریسن-وودباری می‌باشد که در این حالت پیچیدگی از مرتبه $O(n^2)$ خواهد بود. همچنین اگر از تجزیه چولسکی $B_k = L_k D_k L_k^T$ استفاده شود دستگاه دارای $O(n^2)$ عمل می‌باشد. به علاوه فاکتورهای L_{k+1} و D_{k+1} را می‌توان از تخمین به روز شده B_{k+1} در همان تعداد عمل که برای B_{k+1}^{-1} مورد نیاز است محاسبه کرد.[5]

استفاده مهم دیگر از تجزیه چولسکی آنست که خاصیت معین مثبت بودن تخمین ماتریس هسین در خطای گرد کردن از بین نمی‌رود.

فرض کنید که $B = LDL^T$ تجزیه چولسکی B باشد جائیکه $L = (lij)$ ماتریسی پایین مثلثی و $D = diag(d_j)$ ماتریسی قطری است. فرض کنیم $\overline{B} = B \pm vv^T$ بهبود یافته مرتبه اول B باشددر اینصورت

$$\overline{B} = LDL^T \pm vv^T = L(D \pm pp^T)L^T$$

---

[۲] *Powel-Symetric-Broyden*
[۳] *Broyden – Fletcher – Goldfarb – Shanno*

p جواب دستگاه $Lp = v$ می‌باشد و تجزیه چولسکی $D \pm pp^T = \widehat{L}\overline{D}\widehat{L}^T$ با یک رابطه بازگشتی ساده محاسبه می‌شود و خواهیم داشت: $\overline{L} = L\widehat{L}$ و در مورد یک اصلاح مثبت $B = B + vv^T$ بردار p و عناصر $\widehat{L}$ و $\widehat{D}$ در یک روش پایدار عددی با استفاده از تنها $\frac{3n^2}{2}$ عمل محاسبه می‌شوند.

۱۰-۱ بررسی نظری روش

فرض کنید تابع غیرخطی (۱-۱) را داشته باشیم که x^* نقطه بهینه آن باشد یعنی $f(x^*) = 0$. حل کردن معادلات غیرخطی به مساله کمترین مربعات غیرخطی[۴]

$$\min F(x) = \frac{1}{2}\|f(x)\|_2^2 = \frac{1}{2} f(x)^T f(x)$$
$$x \in R^n$$

وابسته است، چون x^* جواب دستگاه می‌باشد.

حال فرض کنید دستگاه معادلات غیر خطی (۱-۱) را به فرم $\min\limits_{x \in R^n} F(x)$ تبدیل کرده باشیم.

بردارگرادیان تابع $F(x) = \frac{1}{2} f^T(x) f(x)$ به صورت $g(x) = \nabla F(x) = J(x)^T f(x)$ می‌باشد که در آن $J(x)$ ماتریس ژاکوبین تابع $f(x)$ بوده و ماتریس هسین آن به صورت $H(x) = \nabla^2 F(x) = J(x)^T J(x)$ تعریف می‌شود.

در محاسبه با روش فوق ریشه مطلوب در مقایسه با روش های نیوتن وشبه نیوتن با نصف تعداد اعمال محاسباتی به دست می‌آید.

روش بهبودمورد نظر، به طور موضعی تابع $F(x)$ را با دستگاه خطی $g(x+k) = g(x) + H(x)S_k$ تقریب می‌زند و به عبارت دیگر مساله بصورت $\min E = \min\|g + HS\|$ حل می‌شود.

اگر تخمین هسین بصورت $B = H$ در نظر گرفته شود:

$$BS = -g$$

خواهیم داشت:

$$B^T BS = -B^T g \qquad (۱۰-۲)$$

حال اگر تخمین هسین B توسط فاکتور QR تجزیه شود: $B = QR$

با جایگذاری عبارت فوق در رابطه (۱۰-۲) خواهیم داشت:

$$(QR)^T (QR)S = -B^T g$$
$$R^T Q^T QRS = -B^T g \qquad (۱۰-۳)$$
$$R^T RS = -B^T g$$

4 Nonliner Least Squares Problems

همان طور که مشاهده می‌شود Q در معادله بالا ظاهر نشد و این یک مزیت روش بهبود می‌باشد که Q زمان اجرای مرحله نیوتن از بین می‌رود و دیگر نیازی به جمع‌آوری Q در حین فرآیند بهبود تجزیه ماتریس در هر تکرار نخواهد بود. این مطلب از میزان محاسبات خواهد کاست که یکی از نتایج آن پایین آمدن خطاست.

طبق فرمول شرمن- موریسن- وودباری، $B_{k+1} = B_k + vv^T$ همگرای مرتبه اول B_k می‌باشد و می‌توان ثابت کرد که $B_{k+1}^T B_{k+1}$ اصلاح شده متقارن مرتبه دوم $B_k^T B_k$ خواهد شد.

حال اگر تجزیه چولسکی را به صورت LDL^T بیان کنیم، خواهیم داشت:

$$
\begin{aligned}
&B = QR \\
&B^T B = LDL^T \quad \rightarrow \quad (QR)^T (QR) = LDL^T \\
&\qquad (R^T Q^T)(QR) = LDL^T \\
&\qquad R^T R = LDL^T = LD^{1/2} D^{1/2} L^T \\
&\qquad R = D^{\frac{1}{2}} L^T \\
&\qquad \rightarrow \quad L = R^T D^{-1/2}
\end{aligned}
$$

۱۱ الگوریتم بهبود یافته

$Tol \in R$ و $x_0 \in R^n$ داده شده است و $f_k = f(x_k)$.

گام۱. $k = 0$.

گام۲. $j_k = j(x_k)$ را محاسبه کنید.

گام۳. مقدار $B_k = j_k^T j_k$ و $G_k = j_k^T f_k$ را محاسبه کنید.

گام۴. تجزیه QR تقریب هسین $B_k = Q_k R_k$ را محاسبه کنید.

گام۵. $D_k = diag(R_{11}^2, ..., R_{nn}^2)$.

گام۶. $L_k = R_k D_k^{-1/2}$.

گام۷. $S_k = -(L_k^T)^{-1} D_k^{-1} L_k^{-1} B_k^T G_k$ را محاسبه کنید.

گام۸. $x_{k+1} = x_k + S_k$.

گام۹. اگر $\|G(x_{k+1})\| \prec Tol$ به گام (۱۳) بروید.

گام۱۰. $Y_k = G(x_{k+1}) - G(x_k)$.

گام۱۱. B_{k+1} را به روز کنید: $B_{k+1} = B_k + \frac{(Y_k - B_k S_k)}{S_k^T B_k S} S_k^T B_k$ [5]

گام۱۲. قرار دهید: $k = k + 1$ و به گام (۴) بروید.

گام۱۳. پایان.

۵ دستور به روز درآوردن ماتریس هسین منسوب به مک کورمیک می باشد.

مثال ۴ جوابهای مثبت دستگاه معادلات غیرخطی زیر را با مقدار اولیه $z_0 = 0.4$ $y_0 = 0.5$ $x_0 = 0.8$ و با الگوریتم شرح داده شده تخمین می‌زنیم. و نتیجه در جدول (۴) قابل مشاهده است.

$$\begin{cases} x^2 + y^2 + z^2 = 1 \\ 2x^2 + y^2 - 4z = 0 \\ 3x^2 - 4y + z^2 = 0 \end{cases}$$

روش PSB		روش **پیشنهادی**	
ROOT	تکرار	*ROOT*	تکرار
$\begin{bmatrix} 0.7851969337 \\ 0.4966113948 \\ 0.3699228297 \end{bmatrix}$	۱۰	$\begin{bmatrix} 0.7851970440 \\ 0.4966115600 \\ 0.3699229703 \end{bmatrix}$	۵

جدول ۴: خطا 10^{-6} در نظر گرفته شده است

۱۲ نتایج ومقایسه روش ها

همان گونه که از مقایسه نتایج به دست آمده از رفتار الگوریتم ها دیده می‌شود:

روش های نیوتن و بریدن نتایج را با دقت مناسب , طول گام کمتر و تقریبا با زمان مساوی ارایه می‌کنند اما روش شبه نیوتن علیرغم مزایایی که دارد نتایج را با همان دقت و طول گام بیشترارایه می‌دهد.

اما اگر بخواهیم آنها را از لحاظ ساختار مقایسه کنیم به نتایج زیر می‌رسیم :

الف- روش نیوتن به طور موضعی همگرای مرتبه دوم است. درهرمرحله محاسبه جواب، n معادله خطی نیاز دارد. دارای الگوریتمی ساده و تئوری قوی است. برای بسیاری از مسائل همگرایی سریعی دارد. همگرایی روش نیوتن تحت تقریب‌های اولیه گوناگون پایدار است اما از دو ضعف جدی از نقطه نظر محاسبات عملی برخوردار است:

۱. حل کردن دستگاه خطی با روش نیوتن از مرتبه $O(n^3)$ است.

۲. زمانی که ژاکوبین به طورتحلیلی موجود نمی‌باشد لازم است تعداد $n^2 + n$ تابع اسکالر در هرتکرار ارزیابی شوند تا تقریب تفاضل متناهی $J(x_k)$ محاسبه شود و این از نظر اقتصادی گران است.

ب- در حالت کلی روش بریدن ماتریس ژاکوبی را بهتر از محاسبه مستقیم آن تقریب می‌زند، بریدن یک تقریب B_k برای $J(x_k)$ به دست می‌آورد که B_{k+1} تقریب $J(x_{k+1})$ از B_k با $O(n^2)$ عمل حسابی در هر تکرار است و ارزیابی f فقط در x_{k+1} , x_k به دست می‌آید و تنها n عمل حسابی درهر تکرار ارزیابی می‌شود. B_{k+1} اصلاح شده مرتبه یک B_k خواهد بود. این روش در بیشتر حالات خوب کار می‌کند ولی دارای این مشکل می باشد که بررسی بد وضع بودن B_k سخت است.

ج- شبه نیوتن روشی بر مبنای روش نیوتن برای حل مجموعه‌ای از معادلات همزمان غیرخطی تعریف می‌کند و ارزیابی‌های $f(x)$ را بیش از آنکه لازم باشد نیاز ندارد بخصوص اگر تابع برداری آن محاسبه سختی داشته باشد. همگرایی روش شبه نیوتن زبرخطی می‌باشد.

مهمترین خاصیت آن است که تقریب ماتریس ژاکوبی ثابت باقی می‌ماند و اگر تقریب ماتریس‌های ژاکوبی B_k به ماتریس ژاکوبی میل کند نرخ همگرایی می‌تواند رشد کند.

اما روش شبه نیوتن دارای ضعف‌هایی نیز می‌باشد از جمله

۱. با استفاده از روش شبه نیوتن همگرایی مرتبه دوم روش نیوتن از بین می‌رود.

۲. روش نیوتن خطای گرد کردن را در تکرارهای متوالی اصلاح خواهد کرد ولی روش شبه نیوتن خود اصلاح نمی‌باشد.

دقت به این نکته نیز لازم است که مقدار بریدن کمترین تغییر B_k نسبت به وضعیت شبه نیوتن است.

زمانی که از فرم پیشنهادی بخش ۱۱ استفاده شد به نتایج زیر دست یافتیم:

۱- در اغلب موارد محاسبه ماتریس هسین از لحاظ محاسبه مستقیم گران تمام می‌شود و این مشکل را می‌توان با تخمین هسین برطرف نمود.

۲- وقتی از مشتق دوم تابع استفاده می‌شود:

الف- محاسبه اعمال حسابی تقریبا به نصف کاهش می‌یابد.

ب- B_k در این حالت بهبود یافته مرتبه دوم B_{k-1} خواهد شد.

۳- در الگوریتم بهبود یافته از تجزیه چولسکی استفاده شد که این عمل احتمال از بین بردن خاصیت معین مثبت بودن ماتریس هسین تخمین زده شده در خطای گرد کردن را منتفی می‌نماید.

۴- در زمان انجام روند نیوتن Q حذف می‌شود و دیگر نیازی به ذخیره‌سازی Q در حین فرآیند بهبود تجزیه ماتریس در هر تکرار نخواهد بود. این مطلب از میزان محاسبات خواهد کاست که یکی از نتایج آن پایین آمدن خطا می‌باشد.

۵- با اجرای این الگوریتم، عناصر $\hat{L}$ و $\hat{D}$ به روش LDL^T در یک روش پایدار عددی با استفاده از تنها $\frac{3n^2}{2}$ عمل محاسبه می‌شوند.

۶- دقت شود که نقطه اولیه به ریشه بقدر کافی نزدیک باشد.

۷- این الگوریتم در مورد ماتریس ژاکوبین نیز قابل اجرا می‌باشد و می‌توان از تجزیه $SQRF$ به جای تجزیه QR استفاده نمود.

۱۳ پیشنهادهای دیگر برای محاسبه ماتریس ژاکوبین

برای هر $x \in R^n$ ماتریس ژاکوبین $J(x)$ را می‌توان به صورت زیر تجزیه کرد:

$$J(x) = M(x)^{-1} N(x) \tag{۱-۱۳}$$

و خواهیم داشت:

$$x_{k+1} = x_k - N_k^{-1} M_k F(x_k)$$

(۱۳-۱) می‌تواند به صورت تجزیه LU و یا تجزیه QR محاسبه شود. اگر ماتریس اسپارس L^{-1} و U را بکار ببریم، به روشی می‌رسیم که در[9] به آن پرداخته شده است.

همچنین برای هر $x \in R^n$ ماتریس ژاکوبین $J(x)$ را می‌توان به صورت زیر بیان کرد:

$$J(x) = C(x) + D(x)$$

که در آن $C(x)$ به راحتی محاسبه می‌شود ولی $D(x)$ این گونه نیست و خواهیم داشت:

$$x_{k+1} = x_k - [C(x_k) + D(x_k)]^{-1} F(x_k)$$

که D_{k+1} تصویر D_k روی زیر فضای آفین $V_{full} = \{D \in R^{n\times n} \mid Ds_k = y_k - C(x_{k+1})s_k\}$ می‌باشد.

روی این موارد در مقاله‌ای که بزودی منتشر خواهد شد کار شده است.

منابع

[1] Demidovich, B.P., Maron, I.A., Mirpublishers, 1981, Computational Mathematics.

[2] Burden, R.L, Faires, J.D, 2001 Numerical Analysis.

[3] Broyden, C.G., 1965, A class of Methods for solving nonlinear simultaneous equations, *Math.Comp.* 19, 577 – 593.

[4] Datta, B. N, 1995, Numerical linear Algebra and Applications.

[5] Wolfe, M.A., 1978, Numerical Methods for Unconstrained Optimization.

[6] More, J.J., Trangenstein, J.A., 1976 on global convergence of Broyden Method, *math Comp.* 30, 523 – 540.

[7] Broyden, C.G., Dennis, J.E., More, J.J., 1973, On the local and super linear convergence of quasi – Newton methods, *J.Inst Math. Appl. V* 12, *PP*. 223 – 245. *MR* 49#6599.

[8] Dennis, J.E, More, J.J, 1974 Charactrization of superlinear Convergence and its application to quasi – Newton methods, *Math. Comp.* 28, 549 – 560.

[9] F.F.Chadee, Sparse quasi-Newton and the continuation problem, T.R.S.O.L. 85 – 8, Department of Operations Research, Stanford University, 1985.