Tentang Stabilitas Gauss-Yordania Eliminasi dengan Pivoting

G. Peters dan J.H. Wilkinson
Laboratorium Fisik Nasional
Teddington, Middlesex, Inggris

Stabilitas algoritma Gauss-Jordan dengan pivot parsial untuk solusi sistem umum persamaan linier umumnya dianggap sebagai tersangka. Hal ini ditunjukkan bahwa dalam banyak hal kecurigaan tidak berdasar, dan secara umum kesalahan absolut dalam solusi ini benar-benar sebanding dengan yang sesuai dengan eliminasi Gaussian dengan pivoting parsial ditambah substitusi kembali. Namun, ketika A dikondisikan dengan buruk, residu yang sesuai dengan larutan Gauss-Jordan seringkali akan jauh lebih besar daripada yang sesuai dengan solusi eliminasi Gaussian.

Kata dan Frasa Kunci: Algoritma Gauss-Jordan, eliminasi Gaussian, substitusi balik, analisis kesalahan mundur, batas untuk kesalahan dalam solusi, terikat untuk residual

Kategori CR: 5.11, 5.14

1. Pendahuluan

Stabilitas numerik penting dari eliminasi Gaussian dengan pivot parsial umumnya ditunjukkan oleh teknik analisis kesalahan mundur [l, 2, 3, 4]. Analisis semacam itu menunjukkan bahwa, ketika sistem n X n diselesaikan pada komputer yang bekerja dalam aritmatika titik Ä mengambang di basis ß dengan mantissa digit-I, solusi yang dihitung adalah solusi yang tepat dari beberapa sistem "tetangga"

(A + E)xc= b.

Istilah "tetangga" digunakan dalam bentuk yang agak longgar

Alamat penulis: Departemen Perdagangan dan Industri, Laboratorium Fisika Nasional, Divisi Analisis numerik dan komputasi, Teddington, Middlesex, Inggris.

rasa. Bahkan, ikatan ditemukan untuk E yang sesuai, yang berbentuk

$\|E\|/\|A\| \leq f(n)g\beta$—t (1.2)

di mana 12 atau norma umumnya digunakan, f(n) adalah fungsi sederhana dari n, dan g adalah faktor "pertumbuhan". Yang terakhir didefinisikan sebagai rasio koefisien modulus maksimum yang terjadi selama eliminasi ke max I aq •,• . Peran yang dimainkan oleh pivoting adalah dalam membatasi kemungkinan pertumbuhan. Meskipun, bahkan dengan pivoting parsial g dapat mencapai nilai $2^{n\,l}$, dalam praktiknya itu umumnya dari urutan kesatuan. Untuk sistem urutan yang lebih besar dari 10, distribusi statistik dari kesalahan pembulatan biasanya memastikan bahwaf(n) > n.

Mari kita menganalisis untuk saat ini konsekuensi dari hasil seperti itu. Misalkan misalnya

$$\| E \|_\infty / \| A \|_\infty \leq n\beta^{-t} \qquad (1.3)$$

Jika kita menulis K = Il A Il , maka disediakan <
0.1 (katakanlah) hubungan (1.1) dan (1.3) memastikan bahwa

$$\| x_c - x \|_\infty / \| x \|_\infty \leq n\beta^{-t}\kappa / (1 - n\beta^{-t}\kappa) \leq (\tfrac{10}{9})n\beta^{-t}\kappa.$$

Oleh karena itu, keakuratan solusi yang dihitung secara langsung tergantung pada K, jumlah kondisi A. Hubungan (1.4) menyiratkan bahwa

$$[1 - (\tfrac{10}{9})n\beta^{-t}\kappa] \| x \|_\infty \leq \| x_c \|_\infty \leq [1 + (\tfrac{10}{9})n\beta^{-t}\kappa \| x \|_\infty \qquad [\text{I} \qquad (1.5)$$

atau, dari asumsi kami bahwa nß-%c 0. l, tentu saja bahwa

$$(\tfrac{8}{9}) \| x \|_\infty \leq \| x_c \|_\infty \leq (\tfrac{10}{9}) \| x \|_\infty. \qquad (1.6)$$

Oleh karena itu Il tentu saja memiliki urutan besarnya yang sama dengan I l x dan ketika nß K jauh lebih kecil dari kesatuan, kedua norma tersebut akan hampir sama. Di sisi lain, vektor residual r yang didefinisikan oleh b — Ax— memenuhi hubungan

$$\| r \|_\infty = \| b - Ax_c \|_\infty = \| Ex_c \|_\infty \leq \| E \|_\infty \| x_c \|_\infty \leq n\beta^{-t} \| A \|_\infty \| x_c \|_\infty. \qquad (1.7)$$

Dengan kata lain, kita memiliki ikatan untuk I l r Il 00 yang hanya bergantung pada ukuran solusi yang dihitung dan bukan pada nomor kondisi A dan oleh karena itu tidak pada keakuratan xc. Kesalahan dalam berkorelasi dengan cara yang memastikan bahwa r biasanya jauh lebih kecil dari yang diharapkan bc ketika K besar. Memang jika kita mengambil kondisi memuaskan vektor acak (1.4), maka untuk yang seperti itu hanya dapat menjamin bahwa

$$\| r \|_\infty = \| b - Ax_c \|_\infty \leq \tfrac{10}{9} \| A \|_\infty n\beta^{-t}\kappa \| x \|$$30. Secara umum perkiraan solusi dengan akurasi yang sama dengan yang diberikan oleh eliminasi Gaussian memberikan residu yang lebih besar dengan faktor K. Bahwa solusi yang dihitung memberikan residu sekecil itu mungkin sangat penting dalam praktiknya. Kita mungkin lebih tertarik pada kedekatan Axc dengan b daripada akurasi absolut xc. Untuk menekankan sifat korelasi yang luar biasa, kami berkomentar bahwa residu yang diberikan oleh solusi yang dihitung memiliki urutan besarnya yang sama dengan yang sesuai dengan solusi yang dibulatkan dengan benar.

Ketika analisis kesalahan mundur dari eliminasi Gauss-Jordan dicoba, ditemukan bahwa seseorang tidak dapat menunjukkan bahwa solusi yang dihitung adalah solusi yang tepat dari beberapa sistem "tetangga" dengan interpretasi yang masuk akal dari kata "tetangga." Kegagalan berasal dari fakta bahwa, dengan Gauss-Jordan, pivoting tidak memberikan kontrol yang memuaskan atas "pertumbuhan." Memang memang sebenarnya tidak lagi benar secara umum bahwa yang dihitung adalah solusi dari sistem tetangga. Untuk alasan ini Gauss-Jordan umumnya dianggap dengan kecurigaan oleh analis numerik. Adalah tujuan dari makalah ini untuk menunjukkan bahwa kecurigaan ini hanya sebagian dibenarkan.

Harus ditekankan bahwa dalam kasus-kasus tertentu seperti ketika A positif pasti atau dominan secara diagonal diketahui bahwa Gauss-Jordan stabil. .

2. Deskripsi Algoritma Gauss-Jordan

Karena algoritma Gauss-Jordan dengan pivoting sudah terkenal, kami akan menjelaskannya hanya secara singkat. Kami menunjukkan sistem aslinya dengan

$$A^{(1)}x = b^{(1)}. \qquad (2.1)$$

Ada n langkah besar. Pada awal langkah rth sistem asli telah digantikan oleh sistem yang setara

$$A^{(r)}x = b^{(r)} \qquad (2.2)$$

di mana $a_{ij}^{(r)} = 0\ (j=1, 2, \ldots, r-1;\ i \neq j)$. Ini berarti bahwa A(r) adalah diagonal sejauh menyangkut kolom r — I pertamanya. Langkah besar rth berlangsung sebagai berikut.

(i) Biarkan max I a, (r)r (Dalam kasus ambiguitas, r dianggap sebagai indeks terkecil.) (ii) Persamaan pertukaran r dan r .

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & \epsilon_2 & 1 & 1 & 1 \\ & & \epsilon_3 & 1 & 1 \\ & & & \epsilon_4 & 1 \\ & & & & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & \epsilon_2^{-1} & \epsilon_2^{-1} & \epsilon_2^{-1} \\ & \epsilon_2 & 1 & 1 & 1 \\ & & \epsilon_3 & 1 & 1 \\ & & & \epsilon_4 & 1 \\ & & & & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & (\epsilon_2\epsilon_3)^{-1} & (\epsilon_2\epsilon_3)^{-1} \\ & \epsilon_2 & 0 & \epsilon_3^{-1} & \epsilon_3^{-1} \\ & & \epsilon_3 & 1 & 1 \\ & & & \epsilon_4 & 1 \\ & & & & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 & (\epsilon_2\epsilon_3\epsilon_4)^{-1} \\ & \epsilon_2 & 0 & 0 & (\epsilon_3\epsilon_4)^{-1} \\ & & \epsilon_3 & 0 & \epsilon_4^{-1} \\ & & & \epsilon_4 & 1 \\ & & & & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ & \epsilon_2 & 0 & 0 & 0 \\ & & \epsilon_3 & 0 & 0 \\ & & & \epsilon_4 & 0 \\ & & & & 1 \end{bmatrix}$$

(iii) Untuk setiap nilai i r, hitung m_{ir} =af r) /a(r r) dan kurangi mq•r kali persamaan r dari persamaan i.

Sistem akhir A x b jelas sedemikian $A^{(n+}$ rupa sehingga diagonal.

Dari pilihan r jelas bahwa m, r < l, (i > r), tetapi untuk i < r tidak ada ikatan seperti itu pada rn"r. Ini berarti bahwa meskipun pertumbuhan elemen di bawah elemen diagonal terbatas seperti dalam eliminasi Gaussian dengan pivoting kolom (memang itu adalah eliminasi Gaussian sejauh menyangkut elemen-elemen ini), pertumbuhan elemen di atas diagonal mungkin secara sewenang-wenang besar. Ini menghalangi kemungkinan analisis kesalahan mundur yang memuaskan yang dianalogikan dengan itu untuk eliminasi Gaussian.

3. Analisis Kesalahan Mundur Standar Gauss-Jordan

Pernyataan kami menunjukkan bahwa bagian dari Gauss-Jordan yang dicurigai adalah produksi nol di atas diagonal. Lebih mudah bagi tujuan kami untuk memikirkan Gauss-Jordan dengan pivoting sebagai terjadi dalam dua tahap yang berbeda: (i) pengurangan ke bentuk segitiga atas oleh algoritma eliminasi Gaussian standar dengan pivot parsial dan (ii) pengurangan lebih lanjut dari sistem segitiga ke sistem diagonal dengan proses eliminasi di mana pivoting dilarang. Pembaca dapat dengan mudah meyakinkan dirinya sendiri bahwa ketika komputasi dilakukan dalam urutan ini kesalahan pembulatan sama dengan dalam prosedur klasik. Perbedaan mendasar antara solusi oleh eliminasi Gaussian dan oleh Gauss-Jordan adalah bahwa pada yang pertama sistem segitiga atas yang dihasilkan diselesaikan

dengan substitusi belakang dan yang terakhir, dengan eliminasi lebih lanjut ke bentuk diagonal. Oleh karena itu kita dapat berkonsentrasi pada stabilitas numerik larutan sistem segitiga atas Ux = c dengan eliminasi.

Pemeriksaan proses ini dengan tujuan untuk melakukan analisis kesalahan mundur segera mengungkapkan diffculty. Ini dapat diekspos melalui contoh sederhana. Pada Gambar I kami menunjukkan pengurangan sistem urutan 5 ke bentuk diagonal. Kami hanya memberikan urutan besarnya jumlah yang dihitung. Dalam matriks segitiga asli diasumsikan bahwa semua elemen memiliki urutan kesatuan kecuali untuk elemen diagonal 1122, 1133, 1144, yang diasumsikan kecil. Kami menulis
-

$\epsilon_i \ (i=2,3,4).$

Akan diamati bahwa, kecuali dalam kasus yang jarang terjadi ketika pembatalan terjadi, pertumbuhan yang cukup besar terjadi dan elemen diturunkan yang sebanding dengan produk timbal balik q. Sekarang dalam analisis kesalahan mundur, gangguan setara yang dihasilkan dari setiap tahap reduksi berbanding lurus dengan ukuran elemen yang muncul dalam matriks yang dikurangi. Oleh karena itu analisis kesalahan mundur menunjukkan bahwa himpunan persamaan diagonal akhir adalah yang akan muncul dari perhitungan yang tepat dengan E di mana terikat untuk I E I diperoleh dari bentuk

$$\beta^{-t}\begin{bmatrix} & 1 & 1 & (\epsilon_2\epsilon_3)^{-1} & (\epsilon_2\epsilon_3\epsilon_4)^{-1} \\ & & 1 & \epsilon_3^{-1} & (\epsilon_3\epsilon_4)^{-1} \\ & 1 & 1 & 1 & \epsilon_4^{-1} \\ 1 & & & 1 & 1 \\ & & & & 1 \end{bmatrix}. \quad €2—1$$

(3.1)

Namun, jika benar bahwa solusi komputasi seakurat yang dapat diharapkan, dengan memperhatikan kondisi U, kita tidak dapat berharap untuk menetapkan ini melalui versi analisis kesalahan mundur yang baru saja kita buat sketsa. Situasi ini sangat kontras dengan holding untuk substitusi belakang dalam sistem segitiga. Di sana mudah untuk menunjukkan bahwa seseorang selalu mendapatkan solusi yang tepat dari beberapa sistem dengan matriks U + E di mana tentu saja I nß I dan karenanya un kecil tidak mempengaruhi matriks E. Namun, fakta bahwa E sekarang sangat besar tidak selalu berarti bahwa solusinya pasti akan sangat buruk. Akan ada banyak set persamaan yang sama sekali berbeda dari Ux = c yang memiliki solusi yang persis sama!

Kami mengamati bahwa gangguan besar dalam diri Anda berada di posisi yang secara khusus terkait dengan posisi q. Mungkinkah gangguan besar terjadi hanya pada posisi-posisi di mana mereka memiliki efek paling kecil? Kami sekarang menunjukkan bahwa ini memang benar untuk contoh yang baru saja kami pertimbangkan. Amati terlebih dahulu bahwa kondisi U setidaknya teratur sehingga bahkan gangguan urutan saja mampu menghasilkan relatif

gangguan dalam solusi ketertiban $(\epsilon_2\epsilon_3\epsilon_4)\beta^{-t}$. Karena kita memiliki gangguan ketertiban (oqeo -lß-t , tampaknya ada bahaya bahwa kita akan mendapatkan gangguan relatif ketertiban $(\epsilon_2\epsilon_3\epsilon_4)^{-2}\beta^{-t}$ dalam solusi. Ketakutan ini terbukti tidak berdasar. Argumen urutan pertama sudah cukup untuk menetapkan ini. Kami punya

(U+E) -l c * — $U^{-1}EU^{-1}c$ = x — (C' E)x. (3.2)

Oleh karena itu kami tertarik pada U'I E, yaitu dalam solusi F dari UF E. Mari kita perhatikan kolom terakhir F. Kami punya

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & \epsilon_2 & 1 & 1 & 1 \\ & & \epsilon_3 & 1 & 1 \\ & & & \epsilon_4 & 1 \\ & & & & 1 \end{bmatrix} \begin{bmatrix} f_{15} \\ f_{25} \\ f_{35} \\ f_{45} \\ f_{55} \end{bmatrix} = \beta^{-t} \begin{bmatrix} (\epsilon_2\epsilon_3\epsilon_4)^{-1} \\ (\epsilon_3\epsilon_4)^{-1} \\ \epsilon_4^{-1} \\ 1 \\ 1 \end{bmatrix}$$

dan segera terlihat bahwa secara umum urutan besarnya f15 dinyatakan oleh hubungan

$$\begin{bmatrix} f_{15} \\ f_{25} \\ f_{35} \\ f_{45} \\ f_{55} \end{bmatrix} \doteq \beta^{-t} \begin{bmatrix} (\epsilon_2\epsilon_3\epsilon_4)^{-1} \\ (\epsilon_2\epsilon_3\epsilon_4)^{-1} \\ (\epsilon_3\epsilon_4)^{-1} \\ \epsilon_4^{-1} \\ 1 \end{bmatrix};$$

tidak ada kuadrat dari semua yang terlibat. Jika gangguan urutan dalam E telah terjadi pada posisi (5,5) atau (5,4), maka f15 dan f25 akan menjadi urutan $(\epsilon_2\epsilon_3\epsilon_4)^{-1}$. Ada sedikit gunanya mencoba mengatur pendekatan ini karena analisis bagian selanjutnya jauh lebih memuaskan, tetapi kami dapat berkomentar di sini tentang satu konsekuensi dari hasil kami. Kami telah menyebutkan sebelumnya bahwa ketika (A + E) x = b maka r - b - Ax = Ex dan ||r II E x Kami tidak dapat memperoleh E kecil dalam analisis kami tentang solusi sistem segitiga dengan eliminasi. Oleh karena itu tampaknya ada bahaya bahwa kesalahan dalam solusi yang dihitung tidak akan berkorelasi sedemikian rupa untuk memberikan residu yang sangat kecil• al.

4. Analisis Kesalahan Terperinci

Stabilitas numerik penting Gauss-Jordan dapat ditetapkan oleh analisis kesalahan mundur yang memiliki tujuan yang agak berbeda dari yang dijelaskan di bagian sebelumnya. Untuk alasan yang sudah diberikan, kita dapat membatasi diri pada pertimbangan sistem segitiga atas.

Mari kita berkonsentrasi sejenak pada operasi yang mengurangi persamaan pertama u11X1 + • • + Lilnxn = Cl ke persamaan yang melibatkan Xl saja. Dijelaskan dalam notasi yang disederhanakan ini dicapai dengan mengurangi dalam $y_2$persamaan waktu suksesi 2, persamaan kali 3, -Pn kali persamaan n dari persamaan pertama. Melupakan kesalahan pembulatan untuk saat yang kita miliki

$$\left.\begin{aligned} &1112 — Y2\,{}^{1}/22 = o, \\ &u_{13} — Y21123\text{ " (Y21123). Y3 }{}^{11}33 = o, \\ &\qquad\qquad \cdots — y_n u_{nn} = 0, \end{aligned}\right\} \quad (4.1)$$

guci — Y3113n — •

dan persamaan turunan akhir adalah

UI IXI = CΓ Y2C2 — Y3C3 — • Yncn. (4.2)

Dalam praktiknya yang dihitung dan ditentukan oleh hubungan

$$
\left.
\begin{aligned}
\bar{y}_2 &= \mathrm{fl}[u_{12}/u_{22}], \\
\bar{y}_3 &= \mathrm{fl}[(u_{13}-\bar{y}_2u_{23})/u_{33}], \\
&\vdots \\
\bar{y}_n &= \mathrm{fl}[(u_{1n}-\bar{y}_2u_{2n}-\bar{y}_3u_{3n}-\cdots-\bar{y}_{n-1}u_{n-1,n})/u_{n,n}], \\
\bar{x}_1 &= \mathrm{fl}[(c_1-\bar{y}_2c_2-\cdots-\bar{y}_nc_n)/u_{11}].
\end{aligned}
\right\} \quad (4.3)
$$

Dengan kata lain, y; dan diturunkan dengan memecahkan sistem segitiga.

L122Y2 = 1112, + [1]/33Y3 [illegible]13,

[illegible]

$$
\begin{aligned}
c_2y_2 + c_3y_3 + \cdots + c_ny_n + u_{11}x_1 &= c_1,
\end{aligned}
\quad (4.4)
$$

dengan proses substitusi ke depan, dan kita tahu dari analisis konvensional back-substitution dalam eliminasi GausSian [3] bahwa proses ini sangat stabil.

Memang nilai-nilai yang dihitung memenuhi persamaan bentuk yang tepat

$$
\begin{aligned}
&u_{22}(1+\epsilon_{22})\bar{y}_2 = u_{12}(1+\epsilon_{12}),\\
&u_{23}(1+\epsilon_{23})\bar{y}_2 + u_{33}(1+\epsilon_{33})\bar{y}_3 = u_{13}(1+\epsilon_{13}),\\
&\quad\vdots\\
&u_{2n}(1+\epsilon_{2n})\bar{y}_2 + u_{3n}(1+\epsilon_{3n})\bar{y}_3 + \cdots +\\
&\qquad + a_{nn}(1+\epsilon_{nn})\bar{y}_n = u_{1n}(1+\epsilon_{1n}),\\
&c_2(1+\epsilon_2)\bar{y}_2 + c_3(1+\epsilon_3)\bar{y}_3 + \cdots + c_{nn}(1+\epsilon_n)\bar{y}_n\\
&\qquad + u_{11}(1+\epsilon_{11})\bar{x}_1 = c_1(1+\epsilon_1).
\end{aligned}
\tag{4.5}
$$

Kami tidak tertarik pada batas yang paling tepat untuk dan Q. Ini akan cukup untuk tujuan kita untuk 0b melayani yang tentunya

$$(1-\epsilon)^n \le 1 + \epsilon_{ij} \le (1+\epsilon)^n,$$
$$(1-\epsilon)^n \le \epsilon_i \le (1+\epsilon)^n,$$

meskipun sebagian besar dan akan memenuhi batas yang lebih ketat; di sini adalah batas untuk kesalahan relatif yang dibuat dalam operasi aritmatika. Lihat, misalnya [31. (Pada komputer biasa yang menggunakan pembulatan, — v}ß ^{1t}). Ini berarti bahwa Sl justru merupakan komponen pertama dari solusi yang tepat . x(^{l}) dari "sistem tetangga."

$$^{(1)})x^{(1)} = c + \delta c^{(1)} \tag{4.7}$$

di mana pasti

$$|\delta U^{(1)}| \le n\epsilon\,|U|, \quad |\delta c^{(1)}| \le n\epsilon\,|c|. \tag{4.8}$$

Beralih sekarang ke komponen kedua kita melihat dengan jenis argumen yang persis sama bahwa itu justru merupakan komponen kedua dari solusi yang tepat dari sistem tetangga.

$$(U+\delta U^{(2)})x^{(2)} = c + \delta c \qquad ^{(2)} \tag{4.9}$$

Matriks öU(^{2}) adalah nol di baris pertamanya. (Persamaan pertama tidak terlibat dalam reduksi persamaan kedua.) Demikian pula komponen pertama dari öc(2) adalah nol. Kami tentu memiliki

$$|\delta U^{(2)}| \le (n-1)\epsilon\,|U|, \quad |\delta c^{(2)}| \le (n-1)\epsilon\,|c| \tag{4.10}$$

dan karenanya fortiori | ö U(^{2}) ne I U I, I öc(^{2}) I I d (4.1 1)

Secara umum xr justru merupakan komponen rth dari larutan x(^{r}) yang tepat dari sistem tetangga

$$^{(r)})x^{(r)} = c + \delta c_{(r)}$$

di mana tentu saja öU(^{r}), dan Dc(r) yang nol di r pertama mereka — I baris, memuaskan

$$_{(r)}| \le n\epsilon\,|U|, \quad |\delta c^{(r)}| \le n\epsilon\,|c|.$$

Perbedaan mendasar antara menyelesaikan sistem segitiga Ux = c oleh Gauss-Jordan dan dengan substitusi belakang adalah bahwa sedangkan untuk yang terakhir seluruh solusi yang dihitung adalah solusi yang tepat dari sistem tetangga tunggal (u + öU) x c -f- öc .

(memang mudah untuk menghindari gangguan öc), dengan yang pertama masing-masing conzponent termasuk dalam solusi yang tepat dari sistem tetangga tetapi ini adalah sistem tetangga yang berbeda untuk masing-masing sistem. Kami sekarang menganalisis konsekuensi dari komentar terakhir ini. Jika x adalah solusi yang tepat dari Ux = c, maka jika

$$^{(r)})x \quad (r) \qquad - c \quad \text{öc} \tag{4.15}$$

Kami punya

$$x^{(r)} \_\_ (U + DU^{O})) \_ c + (U + \text{öU} \quad) \text{öc(r)},$$

$$\_\_ X + + f(i) \text{ (katakanlah), di} \tag{4.16}$$

mana

$$\| \le \left\{ \frac{\| \quad \| \quad \|}{1 - \| U^{-1} \| \, \| \delta U^{(r)} \|} \right\} \text{J ll di I} \tag{4.17}$$

(r)

Ll Si Si

$$\text{ll/(}^{r}\text{) ll} < \quad \text{U I ll ll} \quad (r)\| \tag{4.18}$$

öU(^{r})

Mengingat bahwa c c II S U ll ll x kami menemukan Ux dan karenanya

$$\| e^{(r)} + f^{(r)} \|$$

$$\text{net ll} \quad \text{Dalam I! + ll u} \quad \text{dalam} \quad \text{x} \tag{4.19}$$

$$\text{UI ll}$$

Jika 1 U-i U ll kita menggunakan norma In yang kita miliki

$$\frac{\| x^{(r+1)} - x \|_\infty}{\| x \|_\infty} \le < \quad 2\text{ne} \quad \frac{1 \quad \text{ne ll} \quad \text{ll u}}{} \tag{4.20}$$

= 2126.

Dalam semua kasus ketidaksetaraan ini hanya berlaku di bawah asumsi bahwa < l . Sekarang sejak . xfrr), kami

pasti Punva

$$| \bar{x}_r - x_r | = | x_r^{(r)} - x_r | \leq \max_i | x_i^{(r)} - x_i | = \| x^{(r)} - x \|_\infty \quad (4.21)$$

and hence

$$\| \bar{x} - x \|_\infty = \max | \bar{x}_r - x_r | \leq \max \| x^{(r)} - x \|_\infty \leq \frac{2n\epsilon\kappa}{1 - n\epsilon\kappa} \| x \|_\infty . \quad (4.22)$$

Tabel 1.

Sistem Asli

	MATRIKS		R.H.S.
.826354	432175	613256.614227	.722872
	. 000547.814712.816328		. 154248
	.	915316.814275	.109844
		.982176	.602286

Persamaan I Setelah Reduksi Pertama

.826354	. oooooo	—643.076 -644.352	- 121 . 146

Persamaan I dan 2 Setelah Reduksi Kedua

.826354	. oooooo	72 .2644	—43.9726
	.	000547.0915516	.0564772

Hasil ini persis seperti yang akan kita peroleh memiliki öU dan öc yang sama mengingat semua komponen yang dihitung.

Sekarang ketika back-substitution digunakan untuk memecahkan himpunan segitiga, sudah diketahui (lihat misalnya [3], hlm. 99—107) bahwa solusi yang dihitung seringkali lebih akurat daripada yang diharapkan, dengan memperhatikan ukuran ö u dan öc yang diturunkan oleh analisis backward-error. Namun, ini tidak terlalu penting di sini. Ingatlah bahwa kita terutama tertarik pada solusi sistem persamaan dengan matriks kuadrat penuh, dan solusi dari sistem segitiga hanyalah paruh kedua dari proses tersebut. Dalam beralih dari sistem asli ke sistem segitiga, kesalahan yang sebanding dengan yang sesuai dengan (4,22) di atas sudah akan dibuat.

saya

Sistem Diagonal Akhir

.826354	000000	.		.341074
	.000547	.		.000336315
		.915316	.	.389482
			.982176	.602286

Tabel 11.

Solusi, Kesalahan dan Residu

Kesalahan	Solusi	Kesalahan	Solusi yang benar untuk 6 angka
0.000378. .		0.000000742. .	Gauss-Jordan
—0.000000714. .		—0.000000399. .	Back-substitusi
0.000000855. .		—0.000000855. .	Ulangi
—0.000000038• •		—0.000000038• •	

0,412746	—.000409	0,413503	o. 000348	0.413155
o. 614835	-0,000092	0,614260	.0667	0.614927
—0,425516	0,000001	—0,425516	.000001	Saya 425517
0,613216	0 .000000	0,613216	.000000	0.613216
Sisa		Sisa		

Oleh karena itu akurasi yang agak luar biasa yang sering diperoleh dalam substitusi belakang sedikit menguntungkan kita.

Kita dapat meringkas ini dengan mengatakan bahwa ketika kita memecahkan sistem kuadrat Ax = b dengan eliminasi Gaussian, solusi yang dihitung adalah solusi yang tepat dari beberapa sistem tetangga (A-I-E)x b, dan terikat untuk E tidak melibatkan K. Ketika diselesaikan oleh Gauss-Jordan, solusi yang dihitung bukanlah solusi yang tepat dari sistem tetangga seperti itu tetapi kesalahan I l xe x I l hanya memiliki urutan besarnya yang sama dengan yang sesuai dengan xc, yang merupakan solusi dari sistem semacam itu.

Situasi analog telah didiagnosis dalam kasus inversi matriks oleh eliminasi Gaussian dan substitusi belakang. Ini tidak terjadi bahwa invers X yang dihitung adalah kebalikan dari beberapa (A-BE) di mana E I l memiliki ikatan yang tidak melibatkan K. Memang benar, bagaimanapun, bahwa kolom rth xr adalah kolom rth dari kebalikan yang tepat dari beberapa tetangga (A + Er), tetapi itu adalah Er yang berbeda untuk setiap kolom.

Beralih sekarang ke residual, fakta bahwa itu adalah öU(ᵣ) yang berbeda untuk setiap komponen adalah quitc serius dalam implikasinya, dan residu yang sesuai dengan solusi Gauss-Jordan seringkali lebih besar daripada yang sesuai dengan substitusi belakang oleh faktor orde K. Perhatikan bahwa ini hanya berarti bahwa solusi Gauss-Jordan memberikan residual yang umumnya dari urutan besarnya seseorang secara alami terkait dengan akurasinya; solusi dengan back-substitution memberikan residu yang jauh lebih kecil maka orang akan mengharapkan, dan kinerja ini dicapai hanya karena korelasi khusus dalam kesalahan.

5. Contoh Numerik

Poin-poin yang dibuat di atas diilustrasikan oleh contoh sederhana dari urutan empat yang hanya memiliki satu pivot kecil. Dalam Tabel I kami memberikan langkah-langkah berturut-turut dalam pengurangan Gauss-Jordan. Perhitungan dilakukan dalam aritmatika desimal floatingpoint 6 digit , tetapi untuk pengenalan yang lebih mudah notasi floating-point standar tidak digunakan. Persamaan yang tidak dimodifikasi dalam tahap reduksi apa pun tidak diulang. Amati bahwa pada tahap pertama reduksi, pertumbuhan dengan faktor 1.000 terjadi pada elemen persamaan pertama sebagai akibat dari penggunaan pivot u,22 = .000547. Dalam Tabel Il kami memberikan solusi komputasi yang diperoleh dengan eliminasi Gauss-Jordan dan Gaussian masing-masing, dan sebagai perbandingan kami juga memberikan solusi yang dibulatkan dengan benar. Kesalahan-kesalahan tersebut adalah urutan besarnya yang diharapkan dengan memperhatikan kondisi jumlah matriks segitiga; backsubstitution memberikan kesalahan yang sedikit lebih besar. (Perhatikan bahwa untuk perbandingan yang lebih adil, penggantian kembali dilakukan tanpa akumulasi produk dalam.) Beralih sekarang ke residu, kita melihat bahwa residu pertama yang sesuai dengan Gauss-Jordan jauh lebih besar daripada yang sesuai dengan substitusi belakang. Komponen besar residual muncul dalam persamaan pertama, dan analisis kesalahan mundur dari Bagian 3 memperkirakan hal ini karena untuk contoh ini dalam persamaan pertama kita memiliki komponen besar E.

Pengakuan. Kami ingin mengucapkan terima kasih kepada Profesor G. H. Golub karena telah menarik perhatian kami pada masalah ini dan untuk merangsang diskusi tentang topik tersebut.

Referensi

1. Forsythe, G.E., dan, Moler, C.B. Solusi Komputer o/' Sistem Aljabar Linier. Prentice-Hall, Englewood ClifTs, NJ, 1967.
2. Wilkinson. J.H. Analisis kesalahan metode langsung inversi matriks. J. ACM 8, 3 (Juli 1961), 281A30.
3. Wilkinson, J.H. Kesalahan Pembulatan dalam Proses Aljabar. Toko Alat Tulis Yang Mulia, London; dan Prentice-Hall, Englewood Cliffs, NJ, 1963.
4. Wilkinson, J.H. Masalah Eigenvalue Aljabar. Oxford University Press, London, 1965.

Saya