The solution of a linear system of equations by means of successive approximations

Rudolf Mehmke and Pavel A. Nekrasov

Mat. Sb., v 16, n 3, (1892), pp. 437-459

This correspondence publishes, with detailed proofs, methods of investigating the convergence of Seidel's method from Professor Mehmke. Methods of the same kind belonging to Prof. Nekrasov, cited here in abbreviated form and without proof, are published in the Appendix to Vol. LXIX of the Notes of the Academy of Sciences (St. Petersburg. 1892), in the article "On the question of solving a linear system of equations with a large number of unknowns by successive approximations"; this article did not include only some modifications to the rules of convergence of Seidel's method, indicated below in the letter of Prof. Nekrasov of June 17 (5), 1892. To this correspondence we should add a letter from Professor Mehmke of January 31 (19), 1892, published in Mat. Sbornik earlier under the following title: "On Seidel's method for solving a system of linear equations with a very large number of unknowns through successive approximations". (Mat. Sb., vol. XVI , vol. 2, pp. 342-345).

Nekrasov to Mehmke

Moscow. March 27 (15), 1892.

Dear Colleague,

According to your wish, I am writing to you in Russian, which is a great relief to me.

For my part, I also, like you, was looking for practically applicable rules of convergence of Seidel's method and I have had some success. Assume that this system is

$$\begin{aligned} a_{1,1}x + a_{1,2}y + \cdots + a_{1,n}u + c_1 &= 0 \\ a_{2,1}x + a_{2,2}y + \cdots + a_{2,n}u + c_2 &= 0 \\ \vdots &= \vdots \\ a_{n,1}x + a_{n,2}y + \cdots + a_{n,n}u + c_n &= 0 \end{aligned}$$

In addition, let us use the symbol $|A|$ to denote the absolute value of A.

I succeeded in proving the following series of new rules.

Rule I. For $s = 0, 1, \ldots, n-1$ let:

$$\begin{aligned} q_{s+1} &= \left|\frac{a_{n,n-s}}{a_{n-s,n-s}}\right| q_1 + \left|\frac{a_{n-1,n-s}}{a_{n-s,n-s}}\right| q_2 + \cdots + \left|\frac{a_{n-s+1,n-s}}{a_{n-s,n-s}}\right| q_s + \\ &+ \left|\frac{a_{n-s-1,n-s}}{a_{n-s,n-s}}\right| + \left|\frac{a_{n-s-2,n-s}}{a_{n-s,n-s}}\right| + \cdots + \left|\frac{a_{1,n-s}}{a_{n-s,n-s}}\right|. \end{aligned}$$

If the largest of the values $q_2, q_3, \ldots, q_n$ is smaller than 1, Seidel's method applied to this system must be be convergent.

This rule is inferior to your excellent rule[1] in its simplicity, but it is still applicable in practice and besides, it is broader than your rule. It may happen that your rule is not applicable, while the rule I above is still useful.

Rule II. If for $i = 1, 2, \ldots, n$ the value of $|a_{i,i}|$ is greater than the sum of the magnitudes $|a_{i,1}|$, $|a_{i,2}|, \ldots$, $|a_{i,i-1}|$, $|a_{i,i+1}|, \ldots$, $|a_{i,n}|$, then Seidel's method should be convergent.

This rule is similar to yours and has the same degree of simplicity as your rule.

Rule III. Let for $i = 1, 2, \ldots, n$,

$$\begin{aligned} p_i &= \left|\frac{a_{i,1}}{a_{i,i}}\right| p_1 + \left|\frac{a_{i,2}}{a_{i,i}}\right| p_2 + \cdots + \left|\frac{a_{i,i-1}}{a_{i,i}}\right| p_{i-1} + \\ &+ \left|\frac{a_{i,i+1}}{a_{i,i}}\right| + \left|\frac{a_{i,i+2}}{a_{i,i}}\right| + \cdots + \left|\frac{a_{i,n}}{a_{i,i}}\right|. \end{aligned}$$

[1] Math. Sborn., vol. XVI, n. 2, p. 342.

If the largest of the values $p_2, p_3, \ldots, p_n$ is less than 1, then Seidel's method should be convergent.

If the conditions mentioned in rule III are satisfied, it is easy to write the maximum limits of errors for approximate values $x_m, y_m, \ldots, u_m$ obtained by Seidel's method after m corrections. If by Δ we denote the greatest of

$$|x_m - x_{m-1}|,\ |y_m - y_{m-1}|, \ldots, |u_m - u_{m-1}|,$$

and by P the largest of the values $p_2, p_3, \ldots, p_n$, then we have:

$$|x - x_m| < \frac{p_1\Delta}{1-P},\ |y - y_m| < \frac{p_2\Delta}{1-P}, \ldots, |u - u_m| < \frac{p_n\Delta}{1-P},$$

if $P < 1$.

Higher error limits for approximate quantities are easily obtained even when Rule II applied.

I expect to publish these results in Russian. If this comes to fruition, I will send you a reprint.

Please accept my assurances of complete esteem and faithfulness.

Mehmke to Nekrasov

Darmstadt. April 4 (March 23), 1892

Esteemed colleague,

I have received the prints of my communication which you were kind enough to have made for me as well as your substantial letter of March 27 (15). Receive my sincere thanks for your efforts. Your new rules for the convergence of Seidel's method have surprised me very much.

I have managed to generalize your rules. I write like you the given system in the form

$$\begin{aligned} a_{1,1}x + a_{1,2}y + \cdots + a_{1,n}u + c_1 &= 0 \\ a_{2,1}x + a_{2,2}y + \cdots + a_{2,n}u + c_2 &= 0 \\ \vdots &= \vdots \\ a_{n,1}x + a_{n,2}y + \cdots + a_{n,n}u + c_n &= 0 \end{aligned}$$

just as I denote the absolute value of the quantity A by $|A|$ in the usual way. The generalization of your rule I is now:

Let for $i = 1, 2, 3, \ldots, n-1$,

$$r_{1,i} = -\frac{a_{i,n}}{a_{n,n}},$$

$$r_{2,i} = -r_{1,i}\frac{a_{n,n-1}}{a_{n-1,n-1}} - \varepsilon_{i,n-1}\frac{a_{i,n-1}}{a_{n-1,n-1}},$$

$$r_{3,i} = -r_{1,i}\frac{a_{n,n-2}}{a_{n-2,n-2}} - r_{2,i}\frac{a_{n-1,n-2}}{a_{n-2,n-2}} - \varepsilon_{i,n-2}\frac{a_{i,n-2}}{a_{n-2,n-2}},$$

$$r_{4,i} = -r_{1,i}\frac{a_{n,n-3}}{a_{n-3,n-3}} - r_{2,i}\frac{a_{n-1,n-3}}{a_{n-3,n-3}} - r_{3,i}\frac{a_{n-2,n-3}}{a_{n-3,n-3}} - \varepsilon_{i,n-3}\frac{a_{i,n-3}}{a_{n-3,n-3}},$$

and so on, where $\varepsilon_{i,k} = 1$ if $i < k$, and $\varepsilon_{i,k} = 0$ if $i \geq k$ further let

$$q_h = \sum_{k=1}^{n-1} |r_{h,k}|, \quad h = 1, 2, 3, \ldots, n.$$

If each of the quantities $q_1, q_3, \ldots, q_n$ is less than 1, then Seidel's method is convergent.

If you exchange $a_{i,k}$ with $a_{k,i}$[2] everywhere in this rule, you get the generalization of your rule III.

It can easily happen that your rule I fails while the above rule still applies. I will give the reason for this later. Let us take as an example the system

$$\begin{aligned} 11x + 7y + 7z + c_1 &= 0 \\ 7x + 11y + 7z + c_2 &= 0 \\ 7x + 7y + 11z + c_3 &= 0 \end{aligned}$$

[2]It must read "$a_{i,k}$ with $a_{n-k+1,n-i+1}$" (Note during printing)

[Note from the translator: the matrix is positive definite]

Your rules do not lead to the goal here, since:

$$p_1 = q_1 = \frac{14}{11}, \quad p_2 = q_2 = \frac{175}{121}, \quad p_3 = q_3 = \frac{2303}{1331}.$$

If, on the other hand, the above rule is applied, the following is found

$$q_1 = \frac{14}{11}, \quad q_2 = \frac{77}{121}, \quad q_3 = \frac{931}{1331}.$$

According to this, the sizes of q_2 and q_3 are smaller than 1, consequently there is convergence.

Since your rules contain only the absolute values of the coefficients $a_{i,k}$, these rules can only lead to the goal, if Seidel's method converges when the signs of the coefficients are changed somehow. Such a system is comparable to an absolutely convergent infinite series. The above rule, on the other hand, is also applicable to systems which are, so to speak, only conditionally convergent, i.e. whose coefficients do not tolerate any change of their sign. The example considered above is an example. Because for the system

$$\begin{aligned} 11x - 7y - 7z + c_1 &= 0 \\ -7x + 11y - 7z + c_2 &= 0 \\ -7x - 7y + 11z + c_3 &= 0 \end{aligned}$$

it is easy to prove that Seidel's method diverges.

[Note from the translator: the matrix above is indefinite and one eigenvalue of the iteration matrix is > 1]

Of course, the above rule can also be extended to the cases in which, according to my suggestion, the unknowns are not improved individually, but in groups.

I would be very grateful if you could inform the "Mathematic Society" of the contents of this letter as well.

In the highest esteem. Yours sincerely

Nekrasov to Mehmke

Moscow, April 14 (2), 1892.

Dear Colleague,

I received your interesting letter of April 4 (March 23), and I will convey its contents to the Moscow Mathematical Society, which will certainly want to publish it in its magazine.

Now it is obvious to me that each of us has a way of deriving very general rules for determining convergence of Seidel's method. At the same time, I have every reason to believe that your proofs and mine are different. Would you be able to tell me in writing the essence of your proofs for their publication in Mat. Sbornik? From my side I have prepared my proofs for print and I have already sent them to a St. Petersburg publishing house for publication. Such a mutual publication of our proofs now seems to me perfectly timely.

If you dared to publish the essence of your proofs in our "Mathematical Collection", I would ask you, if possible, to point out what means your method of proof has for judging the speed of convergence of Seidel's method.

As for the generality of the expression of the rules of convergence, it is easy to prove the following general rules (here I am following the notation used in your last letter).

Rule IV. Let

$$\begin{array}{c} \lambda_{1,1}, \\ \lambda_{2,1},\ \lambda_{2,2}, \\ \lambda_{3,1},\ \lambda_{3,2},\ \lambda_{3,3}, \\ \cdots\ \cdots\ \cdots\ \cdots \\ \lambda_{n,1},\ \lambda_{n,2},\ \ldots,\ \lambda_{n,n}, \\ \mu_1,\ \mu_2,\ldots,\mu_n \end{array}$$

be arbitrary quantities, of which $\lambda_{1,1}, \lambda_{2,2}, \ldots, \lambda_{n,n}, \mu_1, \mu_2, \ldots, \mu_n$ are nonzero (and, without limiting the generality can be chosen as $\lambda_{1,1} = \lambda_{2,2} = \cdots = \lambda_{n,n} = \mu_1 = 1$). Let for $k = 1, \ldots, n$, $s = 1, \ldots, k$,

$$b_{k,s} = (\lambda_{k,s}\, a_{n-s+1,n-s-1} + \lambda_{k,s+1}\, a_{n-s+1,n-s} + \cdots + \lambda_{k,k}\, a_{n-s+1,n-k+1})\mu_s,$$

$\varepsilon_{h,s} = 1$ for $h < s$ and $\varepsilon_{h,s} = 0$ for $h \geq s$, and for $k = 1, \ldots, n$, $s = 2, \ldots, k$,

$$\beta_{k,s} = (\lambda_{k,1}\, a_{n-s+1,n}\, \varepsilon_{1,s} + \lambda_{k,2}\, a_{n-s+1,n-1}\, \varepsilon_{2,s} + \cdots + \lambda_{k,k}\, a_{n-s+1,n-k+1}\, \varepsilon_{k,s})\mu_s,$$

$$Q_1 = \left|\frac{\beta_{1,2}}{b_{1,1}}\right| + \left|\frac{\beta_{1,3}}{b_{1,1}}\right| + \cdots + \left|\frac{\beta_{1,n}}{b_{1,1}}\right|,$$

$$Q_k = \left|\frac{\beta_{k,2}}{b_{k,k}}\right| + \left|\frac{\beta_{k,3}}{b_{k,k}}\right| + \cdots + \left|\frac{\beta_{k,n}}{b_{k,k}}\right| + \left|\frac{b_{k,1}}{b_{k,k}}\right| Q_1 + \left|\frac{b_{k,2}}{b_{k,k}}\right| Q_2 + \cdots + \left|\frac{b_{k,k-1}}{b_{k,k}}\right| Q_{k-1}, \; k = 2, 3, \ldots, n.$$

Let Q be the largest of the values $Q_2, Q_3, \ldots, Q_n$. If $Q < 1$, then Seidel's method must be convergent.

Rule V. Let

$$\begin{array}{c} \lambda_{1,1}, \\ \lambda_{2,1}, \; \lambda_{2,2}, \\ \lambda_{3,1}, \; \lambda_{3,2}, \; \lambda_{3,3}, \\ \cdots \; \cdots \; \cdots \; \cdots \\ \lambda_{n,1}, \; \lambda_{n,2}, \ldots, \lambda_{n,n}, \\ \mu_1, \mu_2, \ldots, \mu_n \end{array}$$

be arbitrary quantities, of which $\lambda_{1,1}, \lambda_{2,2}, \ldots, \lambda_{n,n}, \mu_1, \mu_2, \ldots, \mu_n$ are nonzero. Let for $k = 1, \ldots, n$, $s = 1, \ldots, k$,

$$b_{k,s} = (\lambda_{k,s}\, a_{s,s} + \lambda_{k,s+1}\, a_{s+1,s} + \cdots + \lambda_{k,k}\, a_{k,s})\mu_s,$$

$\varepsilon_{h,s} = 1$ for $h < s$ and $\varepsilon_{h,s} = 0$ for $h \geq s$, and for $k = 1, \ldots, n$, $s = 2, \ldots, k$,

$$\beta_{k,s} = (\lambda_{k,1}\, a_{1,s}\, \varepsilon_{1,s} + \lambda_{k,2}\, a_{2,s}\, \varepsilon_{2,s} + \cdots + \lambda_{k,k}\, a_{k,s}\, \varepsilon_{k,s})\mu_s,$$

$$P_1 = \left|\frac{\beta_{1,2}}{b_{1,1}}\right| + \left|\frac{\beta_{1,3}}{b_{1,1}}\right| + \cdots + \left|\frac{\beta_{1,n}}{b_{1,1}}\right|,$$

$$P_k = \left|\frac{\beta_{k,2}}{b_{k,k}}\right| + \left|\frac{\beta_{k,3}}{b_{k,k}}\right| + \cdots + \left|\frac{\beta_{k,n}}{b_{k,k}}\right| + \left|\frac{b_{k,1}}{b_{k,k}}\right| P_1 + \left|\frac{b_{k,2}}{b_{k,k}}\right| P_2 + \cdots + \left|\frac{b_{k,k-1}}{b_{k,k}}\right| P_{k-1}, \; k = 2, 3, \ldots, n.$$

Let P be the largest of the values $P_2, P_3, \ldots, P_n$. If $P < 1$, then Seidel's method must be convergent.

These two general rules, which you can probably easily prove with your own methods, encompass all the rules you and I recently found, including the ones you get with the grouping of equations. The rule you stated in your last letter is derived from the previous rule IV in the case where $\mu_1 = \mu_2 = \cdots = \mu_n = 1$ and the quantities $\lambda_{k,s}$ are chosen such that each of the quantities $b_{h,i}$ with $h < i$ are zero.

In conclusion, let me correct a small inaccuracy in your last letter. You write about the general rule you found: "If you swap $a_{i,k}$ with $a_{k,i}$ everywhere in this rule, you get the generalization of your rule III". It seems to me, here instead of "$a_{i,k}$ and $a_{k,i}$" it should be: "$a_{i,k}$ with $a_{n-k+1,n-i+1}$". This small inaccuracy again gives me reason to guess that your proofs are different from mine. Therefore, the simultaneous publication of yours and my proof is now highly desirable. I look forward to your letter regarding these proofs.

Please accept my assurance of my sincere esteem and perfect faithfulness to you.

Mehmke to Nekrasov

Darmstadt. May 21 (9), 1892.

Dear Colleague,

Receive my best thanks for your kind letter of April 14 (2) this year. Unfortunately, due to lack of time, it was not possible for me to answer it immediately.

In accordance with your request, I am sharing with you today my proofs of the rules for the convergence of Seidel's procedure, which you had established in your letter of March 27 (15) and for which I proposed a somewhat more general version in my letter of April 4 (March 23). In connection with this, I will develop some new convergence criteria.

Let the given system be

$$\begin{aligned} a_{1,1}x + a_{1,2}y + \cdots + a_{1,n}u + c_1 &= 0 \\ a_{2,1}x + a_{2,2}y + \cdots + a_{2,n}u + c_2 &= 0 \\ \vdots &= \vdots \\ a_{n,1}x + a_{n,2}y + \cdots + a_{n,n}u + c_n &= 0 \end{aligned}$$

If one sets in these equations the approximate values $x_i, y_i, z_i, \ldots, u_i$, $(i = 0, 1, 2, \ldots)$ in place of the unknowns $x, y, z, \ldots, u$ so the left-hand sides become $w_1^{(i)}, w_2^{(i)}, \ldots, w_n^{(i)}$ which I will call the inconsistencies (dissent) of the equations in question. While Seidel aimed at reducing the sum of the squares of the inconsistencies by applying improvements to the approximate values of the searched quantities, I have considered another function of the inconsistencies, whose unlimited decrease until disappearance also guarantees the achievement of the goal, namely the sum of the absolute values of the inconsistencies.

Allow me to make an intermediate remark.

By geometrical considerations I have found that for any condition of the system of coefficients a the approximate value of any unknown can always be improved in such a way that the sum of the absolute values of the inconsistencies is not only reduced at all, but by the largest amount, and further, that after applying this improvement, one equation of the system is always exactly fulfilled. But this exactly fulfilled equation is not always the one in which -according to Seidel's expression- the unknown in question is in the diagonal (i.e. at one of the coefficients $a_{1,1}, a_{2,2}, \ldots, a_{n,n}$). The idea suggests itself to modify Seidel's method accordingly, i.e. not to improve any unknown always with the help of the equation which contains this unknown in the diagonal, but to ensure only that with each improvement of an unknown the sum of the absolute values of the inconsistencies is reduced as much as possible. Unfortunately, one does not always reach the goal in this way, because it can easily happen that some individual equations are not used at all and therefore are not affected by the improvements to the extent that the inconsistencies of these equations could be reduced below any amount.

From now on, I want to presuppose that one either proceeds completely according to the prescription of Seidel, or, according to my suggestion, that the approximate values of the unknowns are improved group-wise. If one demands that the sum of the absolute values of the inconsistencies is reduced at each individual step, then one easily finds the convergence rules[3] communicated in my letter of January 31 (19). If, however, one only requires that after a single improvement of the approximate values of all unknowns a decrease of that sum has taken place, then, as I will show shortly, convergence rules of the kind of your rule I result.

It is easy to see that the inconsistencies $w^{(i+1)}$ are linear functions of the inconsistencies $w^{(i)}$, where the coefficients in these linear functions depend only on the order in which the unknowns are introduced and on whether one computes their approximate values individually or in groups and of the values of the coefficients a. If each time with the simultaneous improvement of the approximation values of the k last unknowns from the k last equations, then for $i = 1, 2, 3, \ldots$

$$w_{n-k+1}^{(i)} = 0,\ w_{n-k+2}^{(i)} = 0, \ldots, w_n^{(i)} = 0.$$

We can consequently set (1)

$$\begin{aligned} w_1^{(i+1)} &= \sum_{h=1}^{n-k} \alpha_{1,h}\, w_k^{(i)}, \\ w_2^{(i+1)} &= \sum_{h=1}^{n-k} \alpha_{2,h}\, w_k^{(i)}, \\ \vdots &= \vdots \\ w_{n-k}^{(i+1)} &= \sum_{h=1}^{n-k} \alpha_{n-k,h}\, w_k^{(i)}. \end{aligned}$$

(In the case $i = 0$ the sums on the right-hand side are of course from $= 1$ to $h = n$).

Now, if $|A|$ is, as usual, the absolute value of A,

$$|w_\ell^{(i+1)}| \leq \sum_{h=1}^{n-k} |\alpha_{\ell,h}|\, |w_h^{(i)}|,$$

and therefore when using the abbreviations

$$R_h = \sum_{\ell=1}^{n-k} |\alpha_{\ell,h}|,$$

[3]See Mat. Sborn., vol. XVI, n. 2, p. 342.

$$|w_1^{(i+1)}| + |w_2^{(i+1)}| + \cdots + |w_{n-k}^{(i+1)}| \leq R_1|w_1^{(i)}| + R_2|w_2^{(i)}| + \cdots + R_{n-k}|w_{n-k}^{(i)}|.$$

So, if the conditions

$$R_1 < 1,\ R_2 < 1, \ldots, R_{n-k} < 1$$

are fulfilled, then

$$|w_1^{(i+1)}| + |w_2^{(i+1)}| + \cdots + |w_{n-k}^{(i+1)}| \leq |w_1^{(i)}| + |w_2^{(i)}| + \cdots + |w_{n-k}^{(i)}|,$$

and we have convergence.

One can also set

$$\bar{R}_h = \sum_{\ell=1}^{n-k} |\alpha_{h,\ell}|,$$

then

$$|w_m^{(i+1)}| \leq \bar{R}_m W^{(i)}, \quad (m = 1, 2, \ldots, n-k),$$

where $W^{(i)}$ denotes the largest of $|w_1^{(i)}|$, $|w_2^{(i)}|$, …, $|w_{n-k}^{(i)}|$. So if the conditions

$$\bar{R}_1 < 1,\ \bar{R}_2 < 1, \ldots, \bar{R}_{n-k} < 1$$

are fulfilled, then each of the values $|w_1^{(i+1)}|$, $|w_2^{(i+1)}|$, …, $|w_{n-k}^{(i+1)}|$ will be smaller than the largest among the values $|w_1^{(i)}|$, $|w_2^{(i)}|$, …, $|w_{n-k}^{(i)}|$ and, consequently, convergence occurs.

To find the quantities α, in the given system of equations for $x, y, z, \ldots, u$, initial values are chosen such that the inconsistency of the h-th equation takes the value 1 and inconsistencies of all other equations vanish. If one corrects those (not precisely known) initial values under this assumption according to the underlying method, then the new inconsistencies of the $(n-k)$ first equations immediately yield as many of the quantities α, because for

$$w_1^{(0)} = 0,\ w_2^{(0)} = 0, \ldots, w_{h-1}^{(0)} = 0,\ w_h^{(0)} = 1,$$

$$w_{h+1}^{(0)} = 0, \ldots, w_n^{(0)} = 0,$$

one obtains by virtue of the equations (1),

$$w_1^{(i)} = \alpha_{1,h},\ w_2^{(i)} = \alpha_{2,h}, \ldots, w_{n-k}^{(i)} = \alpha_{n-k,h}.$$

In order to find all α, one has to set h in order equal to $1, 2, 3, \ldots, n-k$.

Let us consider the special case in which the approximate values of the sought quantities are improved individually in the order $x, y, z, \ldots, u$. In this case the quantities α can also be calculated recursively. Let us assume as above that

$$w_1^{(0)} = 0,\ w_2^{(0)} = 0, \ldots, w_{h-1}^{(0)} = 0,\ w_h^{(0)} = 1,$$

$$w_{h+1}^{(0)} = 0, \ldots, w_n^{(0)} = 0.$$

Then only the approximate value of the h-th unknown receives a nonzero improvement $-1/a_{h,h}$ and after applying it the inconsistencies of the n given equations become

$$-\frac{a_{1,h}}{a_{h,h}},\ -\frac{a_{2,h}}{a_{h,h}}, \ldots, -\frac{a_{h-1,h}}{a_{h,h}},\ 0,\ -\frac{a_{h+1,h}}{a_{h,h}}, \ldots, -\frac{a_{n,h}}{a_{h,h}},$$

in turn.

If one now calculates further, i.e. one improves also the approximate value of the $(h+1)$th, $(h+2)$th,..., n-th unknowns, then the amounts added to the $(h-1)$ first of those inconsistencies are obviously equal to the inconsistencies, which would have resulted if the original inconsistencies had been:

$$w_1^{(0)} = 0,\ w_2^{(0)} = 0, \ldots, w_{h-1}^{(0)} = 0,\ w_h^{(0)} = 0,$$

$$w_{h+1}^{(0)} = -\frac{a_{h+1,h}}{a_{h,h}}, \ldots, w_n^{(0)} = -\frac{a_{n,h}}{a_{h,h}},$$

and the same is true of the inconsistencies which the $(n+1-h)$ last equations will finally show. With the help of the equations (1) we get consequently (*)

$$\alpha_{\ell,h} = -\varepsilon_{\ell,h}\frac{a_{\ell,h}}{a_{h,h}} - \alpha_{\ell,h+1}\frac{a_{h+1,h}}{a_{h,h}} - \alpha_{\ell,h+2}\frac{a_{h+2,h}}{a_{h,h}} - \cdots - \alpha_{\ell,n}\frac{a_{n,h}}{a_{h,h}},$$

with $\varepsilon_{\ell,h} = 1$ for $\ell < h$ and $\varepsilon_{\ell,h} = 0$ for $\ell \geq h$. Because

$$\alpha_{\ell,n} = -\frac{a_{\ell,n}}{a_{n,n}}$$

the quantities $\alpha_{\ell,h}$ belonging to the same value of ℓ can be calculated with the help of equation (*) in the order $\alpha_{\ell,n}, \alpha_{\ell,n-1}, \alpha_{\ell,n-2}, \ldots$. The conditions (4), in which now $k = 1$, together with equations (3) and (*) in other notation give your rule I in the further version I propose.

Much sharper convergence criteria than those under (4) and (4') are obtained by the assumption that only after repeated improvement of the approximate values of all unknowns, the sum of the absolute inconsistencies of all equations is forced to fall below the initial amount. By m-times application of the linear substitutions (1) equations of the form

$$w_\ell^{(i+m+1)} = \sum_{h=1}^{n-k} \alpha_{\ell,h}^{(m)} w_h^{(i)} \quad (5)$$

will result, which with the terms

$$R_h^{(m)} = \sum_{\ell=1}^{n-k} |\alpha_{\ell,h}^{(m)}| \quad (6)$$

and

$$\bar{R}_h^{(m)} = \sum_{\ell=1}^{n-k} |\alpha_{h,\ell}^{(m)}| \quad (6')$$

lead to the conditions

$$R_1^{(m)} < 1,\ R_2^{(m)} < 1, \ldots, R_{n-k}^{(m)} < 1, \quad (7)$$

and

$$\bar{R}_1^{(m)} < 1,\ \bar{R}_2^{(m)} < 1, \ldots, \bar{R}_{n-k}^{(m)} < 1, \quad (7')$$

guaranteeing the convergence.

Let us take as an example the system

$$\begin{aligned} x + 0.3y - 0.2z - 4 &= 0 \\ 3x + y - z - 11 &= 0 \\ 2.5x + y + z - 20 &= 0 \end{aligned}$$

whose solution is $x = 2$, $y = 10$, $z = 5$. We have

$$w_1^{(i+1)} = w_1^{(i)} - 0.5w_2^{(i)}, \quad w_2^{(i+1)} = 0.5w_2^{(i)} - w_2^{(i)},$$

$$\alpha_{1,1} = 1,\ \alpha_{1,2} = -0.5,\ \alpha_{2,1} = 0.5,\ \alpha_{2,2} = -1,$$

$$R_1 = 1.5,\ R_2 = 1.5,\ \bar{R}_1 = 1.5, \bar{R}_2 = 1.5.$$

Conditions (4) and (4') are not satisfied, as can be seen. But,

$$w_1^{(i+2)} = 0.75w_1^{(i)}, \quad w_2^{(i+2)} = 0.75w_2^{(i)},$$

i.e.

$$\alpha_{1,1}^{(1)} = 0.75,\ \alpha_{1,2}^{(1)} = 0,\ \alpha_{2,1}^{(1)} = 0,\ \alpha_{2,2}^{(1)} = 0.75,$$

$$R_1 = R_2 = \bar{R}_1 = \bar{R}_2 = 0.75 < 1,$$

Consequently, convergence must occur. This is understandable, because one has

$$|w_1^{(i+2)}| + |w_2^{(i+2)}| = 0.75(|w_1^{(i)}| + |w_2^{(i)}|),$$

i.e. after calculating twice through the whole system the sum of the absolute values of the inconsistencies is only equal to 3/4 of the previous amount.

I think I know how you arrived at your rules II and III. Presumably you start from the equations (4') of your treatise "Determining the unknown by the method of least squares with a very large number of unknowns" in

Mat. Sbornik of the Mathematical Society in Moscow, year 1885, namely (in other notation)

$$\begin{aligned}
a_{1,1}\xi_{i+1} + a_{1,2}\eta_i + a_{1,3}\zeta_i + \cdots + a_{1,n}\nu_i &= 0 \\
a_{2,1}\xi_{i+1} + a_{2,2}\eta_{i+1} + a_{2,3}\zeta_i + \cdots + a_{2,n}\nu_i &= 0 \\
a_{3,1}\xi_{i+1} + a_{3,2}\eta_{i+1} + a_{3,3}\zeta_{i+1} + \cdots + a_{3,n}\nu_i &= 0 \\
\vdots &= \vdots \\
a_{n,1}\xi_{i+1} + a_{n,2}\eta_{i+1} + a_{n,3}\zeta_{i+1} + \cdots + a_{n,n}\nu_{i+1} &= 0
\end{aligned}$$

where $\xi_i, \eta_i, \zeta_i, \ldots, \nu_i$ denote the errors of the approximations $x_i, y_i, z_i, \ldots, u_i$. If one looks for the condition under which the error of the last improved approximation value is definitely smaller, taken in absolute values, than the largest of the errors of the previously improved approximation values, one actually obtains your rule II. If one is satisfied with demanding that each of the errors $\eta_{i+1}, \zeta_{i+1}, \ldots, \nu_{i+1}$ is, in absolute terms, smaller than the largest of the errors $\eta_i, \zeta_i, \ldots, \nu_i$ that is how you arrive at your rule III. Allow me to point out that a rule of similar form is obtained when condition

$$|\eta_{i+1}| + |\zeta_{i+1}| + \cdots + |\nu_{i+1}| < |\eta_i| + |\zeta_i| + \cdots + |\nu_i|$$

is imposed. It relates to your Rule III as Rule (4) of my letter today relates to Rule (4').

Of course, much stricter criteria result, if one merely makes the demand, that for any integer m, which is greater than 1, each of the errors $\eta_{i+m}, \zeta_{i+m}, \ldots, \nu_{i+m}$ taken in absolute values should be smaller than the largest among the errors $\eta_i, \zeta_i, \ldots, \nu_i$ or also, it should be

$$|\eta_{i+m}| + |\zeta_{i+m}| + \cdots + |\nu_{i+m}| < |\eta_i| + |\zeta_i| + \cdots + |\nu_i|.$$

For example, let us look again at the system

$$\begin{aligned}
x + 0.3y - 0.2z - 4 &= 0 \\
3x + y - z - 11 &= 0 \\
2.5x + y + z - 20 &= 0
\end{aligned}$$

Here you find

$$\eta_{i+1} = 0.9\eta_i + 0.4\zeta_i, \quad \zeta_{i+1} = -0.15\eta_i - 0.9\zeta_i.$$

However, since

$$0.9 + 0.4 > 1, \quad 0.15 + 0.9 > 1,$$

your rule III has no force in this case. But,

$$\eta_{i+2} = 0.75\eta_i, \quad \zeta_{i+2} = 0.75\zeta_i,$$

so nevertheless, as we had already seen above, convergence takes place. The law of symmetry you discovered, according to which one can derive another convergence rule by simply substituting each quantity $a_{h,i}$ by the quantity which is symmetric with respect to the transverse diagonal $a_{1,n}, a_{2,n-1}, \ldots, a_{n,1}$ - in my last letter I had written $a_{i,h}$ by mistake - is most interesting. I am very curious about your proof of this.

My proofs for the limits of the errors you have established will hardly differ from yours, which is why I refrain from communicating them. I have not yet succeeded in finding new error limits, however, I have not yet found the time to pay special attention to this subject.

As far as your rules IV and V are concerned, I have not yet been able to form a judgment about their meaning and significance. It would be very desirable to me if you would show the application of these new rules to the example of a system of 3 equations treated in this letter.

With the assurance of the greatest respect and faithfulness.

Moscow. 17 (5) June 1892.

Dear Colleague,

Thank you very much for your letter of May 21 (9), revealing your proofs of the rules of Seidel convergence and intended for publication in Mat. Sbornik.

As for the meaning of rules IV and V in my letter to you from April 14 (2), I must say that a successful choice of arbitrary $\lambda_{h,k}$ and μ_k as part of these rules can make it possible to solve the problem of the Seidel convergence when more simple rules do not lead to the goal. Thus, for the system

$$\begin{aligned} x - 33y - 6z + c_1 &= 0 \\ y + 3z + c_2 &= 0 \\ x + 2y - 100z + c_3 &= 0 \end{aligned}$$

the rules indicated in my letter of March 27 (15), as well as the rules in your letter of March 4 (23), or rules (4) and (4') in your letter of 21 (9) May do not reach their goal. If you apply rule V, using $\mu_1 = \lambda_{1,1} = \lambda_{2,2} = \lambda_{3,3} = 1$, $\lambda_{2,1} = 0$, $\lambda_{3,1} = -1$, $\lambda_{3,2} = -2$, $\mu_2 = 10$, $\mu_3 = \sqrt{11}$, we will have

$$b_{1,1} = 1,\ b_{2,1} = 0,\ b_{2,2} = 10,\ b_{3,1} = b_{3,2} = 0,$$

$$b_{3,3} = -100\sqrt{11},\ b_{1,2} = -330,\ b_{3,3} = 0,$$

$$P_1 = 330 + 6\sqrt{11}, \quad P_2 = P_3 = \sqrt{0.99} < 1,$$

i.e., Seidel's method is convergent. This convergence can be found even with $m = 1$ by means of rule (7) in your letter of May 21 (9).

I have tried with the help of rules IV and V to interpret the system (1)

$$\begin{aligned} x + 0.3y - 0.2z - 4 &= 0 \\ 3x + y - z - 11 &= 0 \\ 2.5x + y + z - 20 &= 0 \end{aligned}$$

which you proposed in your letter of May 21 (9); but I was unsuccessful in my attempt. In order to successfully interpret system (1), I had to modify rules IV and V, by increasing their sensitivity (and the number of computations was by necessity greater).

The suggested modification of rule V can be stated in the following form.

If $x_i, y_i, z_i, \ldots, u_i$ is a system of approximate values, obtained by Seidel's method, then (2)

$$\begin{aligned} a_{1,1}x_{i+1} + a_{1,2}y_i + +a_{1,3}z_i + \cdots + a_{1,n}u_i + c_1 &= 0 \\ a_{2,1}x_{i+1} + a_{2,2}y_{i+1} + z_{2,3}z_i + + \cdots + a_{2,n}u_i + c_2 &= 0 \\ \vdots &= \vdots \\ a_{n,1}x_{i+1} + a_{n,2}y_{i+1} + a_{n,3}z_{i+1} + \cdots + a_{n,n}u_{i+1} + c_n &= 0 \end{aligned}$$

Assume (3)

$$\begin{aligned} x_i &= A_{1,1}X_i, \\ y_i &= A_{2,1}X_i + A_{2,2}Y_i, \\ z_i &= A_{3,1}X_i + A_{3,2}Y_i + A_{3,3}Z_i, \\ \vdots &= \vdots \\ u_i &= A_{n,1}X_i + A_{n,2}Y_i + \cdots + A_{n,n}U_i, \end{aligned}$$

where $A_{k,h}$ are arbitrary coefficients, the choice of which can be used for the best achievement of the goal, and coefficients $A_{1,1}, A_{2,2}, \ldots A_{n,n}$ must be different from zero.

Using equations (2) and (3), it is easy to obtain expressions of $\Delta Y_i, \Delta Z_i, \ldots, \Delta U_i$ from $\Delta Y_{i-1}, \Delta Z_{i-1}, \ldots, \Delta U_{i-1}$ assuming $\Delta M_i = M_{i+1} - M_i$. Let these expressions be:

$$\begin{aligned} \Delta Y_i &= \gamma_{1,1}\Delta Y_{i-1} + \gamma_{1,2}\Delta Z_{i-1} + \cdots + \gamma_{1,n-1}\Delta U_{i-1} \\ \Delta Z_i &= \gamma_{2,1}\Delta Y_{i-1} + \gamma_{2,2}\Delta Z_{i-1} + \cdots + \gamma_{2,n-1}\Delta U_{i-1} \\ \vdots &= \vdots \\ \Delta U_i &= \gamma_{n-1,1}\Delta Y_{i-1} + \gamma_{n-1,2}\Delta Z_{i-1} + \cdots + \gamma_{n-1,n-1}\Delta U_{i-1} \end{aligned}$$

Let us then take

$$\rho_h = \sum_{\ell=1}^{n-1} |\gamma_{h,\ell}|, \quad (h = 1, 2, \ldots, n-1).$$

Seidel's method should be convergent if

$$\rho_1 < 1,\ \rho_2 < 1, \ldots, \rho_{n-1} < 1.$$

Using your methods as indicated in the letter of May 21 (9), 2011, and based on the consideration of the sum of $|Y_{i+1} - Y_i| + |Z_{i+1} - Z_i| + \cdots + |U_{i+1} - U_i|$, instead of which you can also look at the sum of

$$|\Delta Y_i| + |\Delta Z_i| + \cdots + |\Delta U_i|,$$

We will also conclude that Seidel's method should be convergent if

$$\bar{\rho}_1 < 1,\ \bar{\rho}_2 < 1, \ldots, \bar{\rho}_{n-1} < 1,$$

with

$$\bar{\rho}_h = \sum_{\ell=1}^{n-1} |\gamma_{\ell,h}|, \quad (h = 1, 2, \ldots, n-1).$$

If further in the expressions ρ_h and p$\bar{\rho}_h$, which depend on the coefficients of the system (2), if we replace the coefficients $a_{i,k}$ with coefficients $a_{n+1-k,n+1-i}$, then the two rules (5) and (6) mentioned above will become two new rules, the first of which will be a modified version of rule IV mentioned in my letter of April 14 (2).

Finally, these four rules are modified by your method, based on looking at expressions of $\Delta Y_{i+m}, \Delta Z_{i+m}, \ldots, \Delta U_{i+m}$ from $\Delta Y_{i-1}, \Delta Z_{i-1}, \ldots, \Delta U_{i-1}$. I do not intend to resort to them when discussing system (1).

Of all the rules listed above, rules (5) and (6) have the peculiarity that they easily lead to expressions of better limits of errors of approximations $x_i, y_i, z_i, \ldots, u_i$.

Now applying rule (5) or (6) to the system (1), let us assume that $A_{1,1} = 1$, $A_{2,1} = 0$, $A_{2,2} = 4$, $A_{3,1} = 0$, $A_{3,2} = -9$, $A_{3,3} = 5\sqrt{3}$. Under these conditions we find:

$$\Delta Y_i = \sqrt{0.75}\Delta Z_{i-1}, \quad \Delta Z_i = \sqrt{0.75}\Delta Y_{i-1}.$$

Consequently, $\rho_1 = \rho_2 = \bar{\rho}_1 = \bar{\rho}_2 = \sqrt{0.75} < 1$, i.e., Seidel's method is convergent.

It goes without saying that practical calculators will prefer the simplest rules given by you and me. But the above general rules for the convergence of Seidel's method in different cases can be applied in the hands of a skilled calculator.

Please accept, dear colleague, the assurance in my deep respect and sincere faithfulness.