Nuclear Instruments and Methods 216 (1983) 177–181
North-Holland Publishing Company

METHOD OF CONVERGENT WEIGHTS – AN ITERATIVE PROCEDURE FOR SOLVING FREDHOLM'S INTEGRAL EQUATIONS OF THE FIRST KIND

A. KONDOR +

Laboratory of Nuclear Problems, Joint Institute for Nuclear Research, 101000 Moscow, USSR

Received 10 January 1983

A method has been developed for solving Fredholm's integral equations of the first kind for the case of positive kernel function and right hand side. The iterative solution seems to be satisfactorily stable against perturbations of a statistical nature. This feature and the basic principle of the iteration itself come from including in the algorithm the so called acceptance function related to the kernel of the integral operator. The main characteristic of the method – to iterate via smooth functions – becomes advantageous compared to standard methods in those cases of integral equations with smooth solutions for which both the kernel function and the right hand side function are strongly nonsmooth, which is often the case for the solution of the smearing problems of the high energy physics experiments.

1. Unfolding problems and integral equations

Unfolding problems arise in measuring the density $u(x)$ of some quantity x when only the distorted (smeared) distribution of x can be observed. x may be one dimensional or higher dimensional and x may be a random variable or a certain physical quantity which can be described by a density function $u(x)$. Since the principles of the problems which are considered here are always the same independently of the dimension of x, in the following, for convenience, x is always considered one dimensional and taking values between 0 and 1.

A common unfolding problem is the following. Let x be a random variable with probability density function $u(x)$. In a great number of the cases a real measurement of x does not yield an ideal sample of the x-values from the density $u(x)$ but one has distorted values y with a density function $v(y)$ where the distortion is caused by some measurement noise w which is independent of x and which has the probability density $p(w)$. This situation leads to the well-known problem of the solution of the convolution integral equation of the first kind.

Very often, however, the independent noise model is not general enough such that more complicated equations than the convolution one have to be solved. First of all the measurement distributions can depend on x such that p becomes a function of two variables and the equation relating the distributions u, v and 8ip generalises to the Fredholm's integral equation of the first kind:

$$v(y)=\int_0^1 p(x,y)\,u(x)\,dx, \qquad 0\leqslant y\leqslant 1 \tag{1}$$

where

$$p(x,y)\geqslant 0 \qquad \text{for } 0\leqslant x,y\leqslant 1.$$

Secondly, it can happen that no y-value is measured even if an x-value occurred. This fact may be expressed by the function

$$A(x)=\int_0^1 p(x,y)\,\mathrm{d}y, \qquad 0\leqslant x\leqslant 1 \tag{2}$$

for this, in this case, the inequality

$$0\leqslant A(x)\leqslant 1, \qquad 0\leqslant x\leqslant 1 \tag{3}$$

holds. The function A is called the acceptance function of the experimental device represented in our case by $p(x,y)$. Suppose that

$$A(x)<1 \tag{4}$$

for all the x in some interval in [0,1]. It then follows from (1) to (4) that

$$\int_0^1 v(y)\,\mathrm{d}y=\int_0^1\int_0^1 p(x,y)\,u(x)\,\mathrm{d}x\,\mathrm{d}y$$
$$=\int_0^1 A(x)\,u(x)\,\mathrm{d}x<1 \tag{5}$$

if it is assumed that the order of integration in the double integral in (5) can be changed. The occurrence of this inequality also causes problems in practice.

Two more problems can be considered. The density v can only be measured with some error and often the function p is not exactly known. For instance, in high energy physics, either p can only be estimated by Monte Carlo simulation or due to computer cost, for an assumed u_0 the corresponding v_0 can be estimated only. Another well-known problem comes from the fact that the problem of solving (1) for given p and v is an

+ Permanent address: Central Research Institute for Physics, 1525 Budapest 114, P.O.B. 49, Hungary.

0167-5087/83/0000–0000/$03.00

ill-posed mathematical problem. That means, for small changes in v large changes in u may be caused. Many different proposals to solve eq. (1) have been made ranging from series solutions over least square solutions to regularization. Since there is no unique method which solves the problem in general, sometimes different methods may have to be tried. This implies that it is important to have different methods available.

In this paper an iterative method is proposed whose convergence properties have not yet been proved mathematically. However, numerous cases of a limited class of eqs. (1) with measurement noise on v and p have shown that good results may be obtained in practice.

In sect. 2 the class of eqs. (1) is given to which the iterative method is applied together with the iteration scheme. In sect. 3 numerical examples are presented.

2. The method of convergent weights

In order to construct an iteration for solving eq. (1) the following assumptions are made to define the class of equations to which the iteration is applied.

(a) The kernel function $p(x, y)$ is non-negative and continuous on $[0, 1]\times[0, 1]$ except possibly isolated integrable singularities.

(b) The acceptance function $A(x)$ is positive on [0,1]

$$0 < A(x) = \int_0^1 p(x, y)\,\mathrm{d}y, \quad 0 \leqslant x \leqslant 1. \tag{6}$$

(c) $$g(y) = \int_0^1 p(x, y) f(x)\,\mathrm{d}x > 0, \qquad 0 \leqslant y \leqslant 1 \tag{7}$$

for every integrable positive function $f(x)$.

(d) For a given $v = v(y)$ which is assumed to be positive and continuous on [0,1] the integral equation

$$v(y) = \int_0^1 p(x, y) u(x)\,\mathrm{d}x, \qquad 0 \leqslant y \leqslant 1 \tag{8}$$

has a unique continuous solution $u = u(x)$ for $0 \leqslant x \leqslant 1$.

(e)

$$\int_0^1 \int_0^1 p(x, y) u(x)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^1 p(x, y) u(x)\,\mathrm{d}y\,\mathrm{d}x. \tag{9}$$

Before constructing the iterative scheme a consequence of the above assumptions will be discussed. Suppose one has a series of iterative approximations $\{u^{(j)}(x)\}_0^\infty$ of the exact solution $u(x)$. Using Eq. (7) one can calculate $v^{(j)}(y)$ the iterative right hand side

$$v^{(j)}(y) = \int_0^1 p(x, y) u^{(j)}(x)\,\mathrm{d}x, \qquad 0 \leqslant y \leqslant 1. \tag{10}$$

Introducing $\delta^{(j)}(x)$, the difference of $u(x)$ and $u^{(j)}(x)$, the following equation can be found

$$\left[1 - \frac{1}{r^{(j)}(y)}\right] v(y) = \int_0^1 p(x, y)\,\delta^{(j)}(x)\,\mathrm{d}x, \qquad 0 \leqslant y \leqslant 1 \tag{11}$$

where

$$\delta^{(j)}(x) = u(x) - u^{(j)}(x) \quad \text{and} \quad r^{(j)}(y) = \frac{v(y)}{v^{(j)}(y)}.$$

Eq. (11) proves that constructing a convergent to the unity series $\{r^{(j)}\}_0^\infty$ under the assumptions (c) and (d) one will have a convergent series of the corresponding $\{u^{(j)}\}_0^\infty$ with the limit $u^{(\infty)} = u(x)$ being the solution of eq. (8).

Forming such an iteration at first we transform the original equation to one having a normalized kernel. It is easy to do via introducing the so called importance function

$$z(x) = A(x) u(x). \tag{12}$$

It then follows from (1), (2) and (6) that

$$v(y) = \int_0^1 q(x, y) z(x)\,\mathrm{d}x, \qquad 0 \leqslant y \leqslant 1 \tag{13}$$

where

$$q(x, y) = p(x, y)/A(x)$$

and q(x, y) is normalized to the unity. Eq. (13) has the property of normalization

$$\int_0^1 v(y)\,\mathrm{d}y = \int_0^1 \int_0^1 q(x, y) z(x)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 z(x)\,\mathrm{d}x. \tag{14}$$

This normalization may have a regularization effect on the iterative solution, so eq. (14) is used as a motivation for requesting to satisfy eq. (14) by all the iterants. This requirement implies the following form of the iterational equation. Choosing a starting function $u^{(0)}(x)$ one can calculate $r^{(0)}(y)$ as

$$v^{(0)}(y) = v(y)/v^{(0)}(y), \qquad 0 \leqslant y \leqslant 1$$

and $z^{(0)}(x)$ as

$$z^{(0)}(x) = A(x) u^{(0)}(x), \qquad 0 \leqslant x \leqslant 1$$

and then the further iterants $z^{(j)}(x)$ will be given by

$$z^{(j+1)}(x) = \int_0^1 r^{(j)}(y) q(x, y)\,\mathrm{d}y \cdot z^{(j)}(x), \qquad 0 \leqslant x \leqslant 1, \tag{15}$$

where equations (7)–(13) were used and $u^{(0)}(x)$ was supposed to be arbitrary but satisfying assumptions (a)–(d).

Eq. (15) has two characteristic features. All the $z^{(j)}(x)$ satisfy the normalization requested for the exact solution $z(x)$ by eq. (14) and the definition of $q(x, y)$ in (13) guarantees the convergence of $z^{(j)}(x)$ in the cases when $r^{(j)}(y) \to 1$. The reason for choosing the

name the Method of Convergent Weights (MCW) for this iteration was the central role of the convergence of the series $\{r^{(j)}\}_0^\infty$ to unity. The ratio $r^{(j)}$ is called a weight because it guarantees as a weighting function the normalization (eq. (14)) until the series of $\{Z^{(j)}\}_0^\infty$ is not converged.

The basic equation of the iteration can be transformed to

$$u^{(j+1)}(x) = \int_0^1 r^{(j)}(y)\, q(x,y)\,\mathrm{d}y \cdot u^{(j)}(x),$$

$$0 \leqslant x \leqslant 1, \tag{16}$$

which is equivalent to

$$u^{(j+1)}(x) = [1 + \mathcal{F}^{(j)}]\, u^{(j)}(x), \tag{17}$$

where

$$\mathcal{F}^{(j)} = \int_0^1 \frac{v(y) - v^{(j)}(y)}{v^{(j)}(y)}\, q(x,y)\,\mathrm{d}y,$$

One can see that the multiplicative operator $\mathcal{F}^{(j)}$ depends on its arguments via an integral of the relative error of the iterative right hand side

$$\frac{Bu - Bu^{(j)}}{Bu}; \qquad Bf = \int_0^1 p(x,y) f(x)\,\mathrm{d}x.$$

Eqs. (16) and (17) are equivalent in the case of exact values but in their numerical realization eq. (17) has advantages giving higher stability against rounding errors and those of the discretization of the integrals.

3. Test cases

Two test cases will be presented to demonstrate the convergence and the stability of the MCW. Both cases are created by analogy with real physical experiments.

3.1. Case 1

In this test case we demonstrate the solution of the one-dimensional smearing problem by the example of a purified numerical model of a real experimental problem [1]. The problems arising from the finite experimental resolution and the imperfect acceptance of a real experiment we model by a Fredholm's integral equation of the first kind:

$$v_t(y) = \int_a^b p_t(x,y)\, u_t(x)\,\mathrm{d}x, \tag{T.1}$$

where a = 4, b = 24, $u_t(x) = 1600/x^2$,

$$p_t(x,y) = \Big[A - B(5-x)^2 - \sum_{k=1}^{3} C_k \exp\{-D_k(x-x_k)^2\}\Big] \times \exp\{-E(x-y)^2\} F,$$

$$A = 1.2, \qquad B = \begin{cases} 1 & \text{for } x < 5, \\ 0.0023 & \text{for others,} \end{cases}$$

$$C_k = 0.4, 0.3, 0.2 \qquad D_k = 5, 5, 1, \qquad x_k = 7, 9, 13$$

$$E = 10, \qquad F = \begin{cases} 1 & \text{for } |x-y| < 0.08x + 0.88, \\ 0 & \text{for others,} \end{cases}$$

and the right hand side $v_t(y)$ is given by straightforward integration (see fig. 1).

We have chosen 50 equidistant coarse mesh-points for discretizing of functions of one variable. The pseudo-experimental errors of the RHS we obtained from the following expression:

$$\tilde{v}_i = v_t(x_i)(1 + \Theta_i p); \qquad i = 1, 2, \ldots, 50, \tag{T.2}$$

where Θ_i are equally distributed random numbers from the interval $[-1, 1]$ and $p = 0.05$.

The possible errors of the kernel function we have modeled in the following manner:

$$p_{ik} = p_t(x_i, y_k)(1 + \Theta_{ik}\tilde{p}); \qquad i, k = 1, 2, \ldots, 50, \tag{T.3}$$

where $\tilde{p} = \max\{0.5, (1 + p\sqrt{|i-k|})\}$.

The discretization of eq. (T.1) has been carried out using Simpson's quadrature with 20 fine mesh-intervals in each coarse one.

The resulting quadratic system of linear equations has been solved by using the MCW and the standard Seidel's method, the latter used for comparison.

For the purpose of illustrating the application of the MCW we tested the convergence in the ideal case, i.e. with errorless $p_t(x,y)$ and $v_t(y)$ but with the errors of the discretization (fig. 1). We started with a trial function $u^{(0)}$ = constant. The relative error of the iterant was found to converge to within $\pm 0.5\%$ in 7 iterations. The

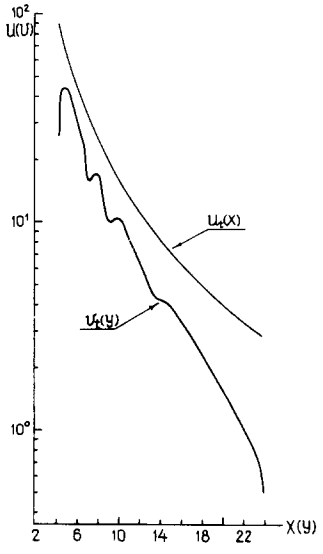

Fig. 1. Functions $u_t(x)$ and $v_t(y)$ for Case 1.

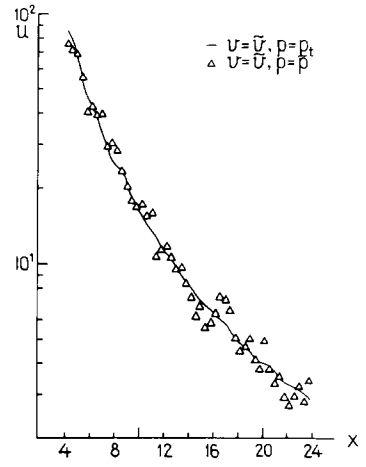

Fig. 2. Fluctuations in the iterated solution in the case of $\pm 5\%$ error in the RHS and ± 5–50% error in the kernel function.

stability of the iteration has been proved in 1000 iteration steps in which case the iterant remained within the $\pm 0.5\%$ error corridor.

Modeling the experimental errors as formulated above we obtained the following results. In the case of $\tilde{v}$ and p_t the result from 1 iteration is shown in fig. 2 by a full line, while the case of $\tilde{v}$ and $\tilde{p}$ is illustrated by triangles. The parallel computation by the Seidel's iteration leads to a solution with fluctuations in a range of $\pm 100\%$ for the case of $\tilde{v}$ and $\tilde{p}$.

The stability of the MCW can be improved by two transformations of the basic equations both having smoothing character. At first one can transform the discretized system of linear equations to the normalized one multiplying by the transposed matrix of the system. As a second step one can apply a proper power function

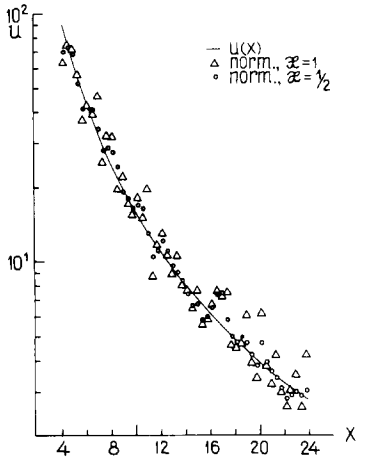

Fig. 3. Changes in the stability of the iterated solution applying smoothing transformations to the kernel function.

(with an exponent $0 < \kappa < 1$) for substituting $r^{(j)}(y)$ by $[r^{(j)}(y)]^{\kappa}$ in eq. (16). Having slowed down the iteration one has to choose a more realistic trial function than $u^{(0)} = \text{const}$.

The change in the stability of the MCW has been investigated in 7 iterations choosing

$$u^{(0)}(x) = 1100/x^2\,[1 + 2\sin(0.2x - 0.8)],$$

which differs from $u_t(x)$ in a range of ± 20–100%. The smoothing properties of the above transformations for $\kappa = \frac{1}{2}$ are shown in fig. 3.

3.2. Case 2

For the second test case we have chosen a typical illposed problem. Our test case is based on that of the monography of Tikhonov and Arsenin [2].

By analogy with the unfolding procedure of radiation spectra, the following mathematical experiment can be constructed. Suppose that the physical process under discussion led to solving the following integral equation:

$$v_t(y) = \int_a^b p_t(x, y)\, u_t(x)\,\mathrm{d}x, \qquad \text{(T.4)}$$

where $a = 0, \quad b = 10,$

$$u_t(x) = \left[1 - \exp\{-A(b-x)^2\}\right] + B\sin\left(\frac{\pi}{b}\cdot x\right) + \sum_{k=1}^{2} C_k \exp\{-D_k(x - x_k)^2\}$$

$$A = 2 \qquad B = 0.5$$

$$C_k = 6, 6 \qquad D_k = 3, 4 \qquad x_k = 3, 7.$$

$$p_t(x, y) = \left(1 - \frac{x}{y}\right)\gamma(y - x); \quad \gamma(z) = \begin{cases} 1 & \text{for } z > 0 \\ 0 & \text{for others} \end{cases}$$

and $v_t(y)$ is given by straightforward integration (see fig. 4).

The discretization has been provided as in Case 1.

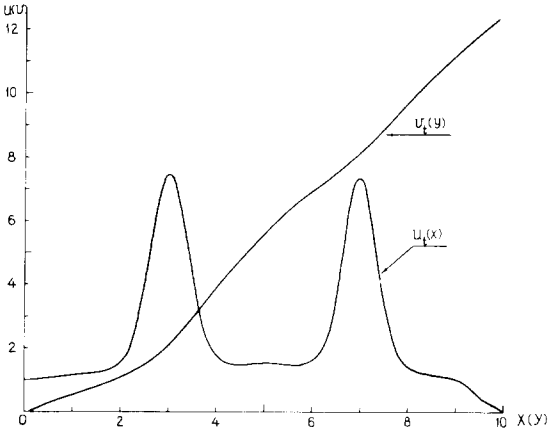

Fig. 4. Functions $u_t(x)$ and $v_t(y)$ for Case 2.

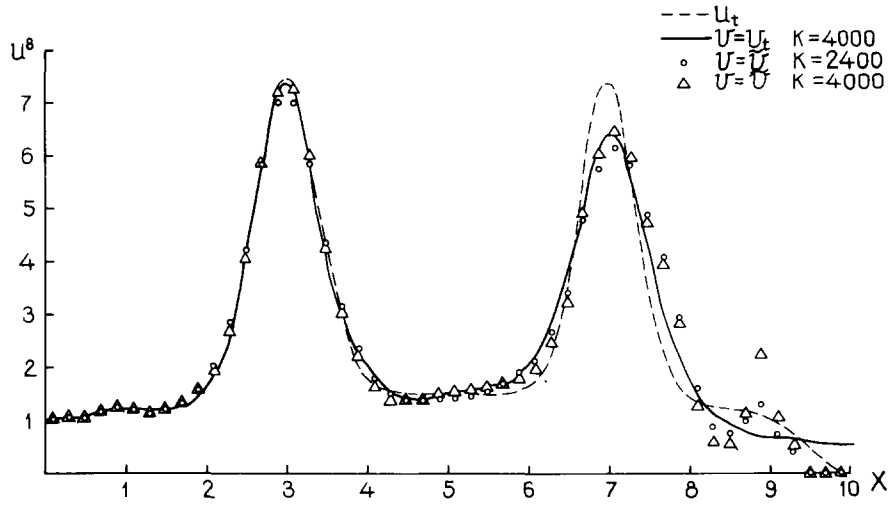

Fig. 5. Comparison of fluctuations excited by ±5% error of the RHS for large numbers of iterations.

The errors of the RHS have been modeled by the formula given in [2]:

$$\tilde{v}_i = v_t(y_i)\left(1 + \Theta_i\sqrt{\frac{3(b-a)}{b^3-a^3}}\,p\right), \qquad i = 1, 2, \ldots, 50,$$

where Θ_i are random numbers from the interval $[-1, 1]$ and $p = 0.05$

Due to the integral in eq. (16) the convergence of the MCW is extremely slow for this test case which give good possibilities for investigating the numerical stability of the iteration. For the ideal case ($p = 0$) the result of 4000 steps is presented in fig. 5 in comparison with the exact solution. Choosing $p = 0.05$ one can see (fig. 5) that the characteristic fluctuations of $u^{(k)}$ appear late enough to allow real peaks to be distinguished from ones arising due to errors of the RHS.

4. Conclusions

The use of the acceptance and the importance functions related to the kernel function of Fredholm's integral equations of the first kind has led to the construction of an iterative procedure which seems to be satisfactorily stable against the errors of the RHS and the kernel function. The convergence is controlled by the integral of the relative error of the iterated RHS that restricts the iteration to be recommended for solving integral equations with nonsmooth RHS.

The effect of the transformation to the normal equation as well as that of the introduction of a power function of the RHS' ratio [in eq. (16)] is to reduce the instability of the MCW due to the errors in the RHS and the kernel function. Application of an exponent in the power function far from the unity makes the convergence slow down, so in such a case one has to choose a more realistic trial function than $u^{(0)} = \text{const}$.

It is hoped that improvements in the direction of regularization of the basic equation will permit the method to be applied into a wide range of the ill-posed problems.

The author would like to thank Drs. B. Schorr (CERN) and G. Emelyanenko (JINR) for helpful discussions.

References

[1] D. Bollini et al., Phys. Rev. Lett. 104B (1981) 403.
[2] A.N. Tikhonov and V. Ya. Arsenin, "Metody resheniya nekorrektnyh zadach" (Methods for Solution of ill-posed Problems), (Nauka, Moscow 1979).