IJST (2014) 38A2: 123-132
Iranian Journal of Science & Technology

http://ijsts.shirazu.ac.ir

Chain least squares method and ill-posed problems

E. Babolian[1], A. Abdollahi[1]* and S. Shahmorad[2]

[1]*Department of Mathematics, Science and Research Branch, Islamic Azad University, P. O. Box 775-14515, Tehran, Iran*
[2]*Faculty of Mathematical Science, University of Tabriz, P. O. Box 51664-16471, Tabriz, Iran*
E-mail: a.abdollahi@srbiau.ac.ir

Abstract

The main purpose of this article is to increase the efficiency of the least squares method in numerical solution of ill-posed functional and physical equations. Determining the least squares of a given function in an arbitrary set is often an ill-posed problem. In this article, by defining artificial constraint and using Lagrange multipliers method, the attempt is to turn n-dimensional least squares problems into $(n-1)$ ones, in a way that the condition number of the corresponding system with $(n-1)$-dimensional problem will be low. At first, the new method is introduced for 2 and 3-term basis, then the presented method is generalized for n-term basis. Finally, the numerical solution of some ill-posed problems like Fredholm integral equations of the first kind and singularly perturbed linear Fredholm integral equations of the second kind are approximated by chain least squares method. Numerical comparisons indicate that the chain least squares method yields accurate and stable approximations in many cases.

Keywords: Lagrange multipliers method; chain least squares; ill-posed problem; integral equations

1. Introduction

The least squares method is used in numerical solution of many physical and engineering problems (Ching and Suh-Yuh, 2002; Alexander and George, 1990; Jinming, 2012; King and Krueger, 2003; Jagalur-Mohan et al., 2013; Laeli and Maalek, 2012; Aksan et al., 2006; Jannike and Hugo, 2012). If this method is applied in an orthogonal basis, the unknown coefficients will be easily computed. The main problem occurs when it is used in a non-orthogonal basis. In this case, determining unknown coefficients leads to solving an ill-posed linear system of equations with large condition number (Datta, 2010; Kincaid and Ward, 2002). This kind of least squares problem occurs in numerical solution of ill-posed functional equations (Delves and Mohamed, 1985; Nashed, 1976). For instance, consider the solution of the following ill-posed functional equation.

$$Ax = f, \quad A: L^2[a,b] \to L^2[a,b], \tag{1}$$

where A is a linear operator and f is a known function. Let $\{p_i\}_{i=0}^{\infty}$ be a basis for $L^2[a,b]$. Then we can approximate the unknown solution x by

$$x_n(t) = \sum_{i=0}^{n} a_i p_i(t). \tag{2}$$

By substituting (2) into (1), we find that

$$\sum_{i=0}^{n} a_i q_i(s) = f(s) + r_n(s), \quad a \le s \le b,$$

where

$$q_i(s) = (Ap_i)(s), \quad i = 0, \dots, n,$$

and $r_n(s)$ is the residual function. In least squares method for determining the unknown coefficients $\{a_i\}_{i=0}^{n}$, we should solve the minimization problem (Datta, 2010).

$$\min_{a_0,\dots,a_n} e(a_0, \dots, a_n),$$

where

$$e(a_0, \dots, a_n) = \int_a^b r_n^2(s)ds = \int_a^b \left[\sum_{i=0}^{n} a_i q_i(s) - f(s)\right]^2 ds$$

In fact, the purpose is to find the least squares approximation of the function f in the set $\{q_i\}_{i=0}^{n}$ although this set is predetermined and is not an orthogonal set. In this article, the ill-conditioning of the least squares problem is postponed to achieve accurate approximations by presenting a chain rule based on the Lagrange multipliers method. We call this approach a chain least squares method. The outline of the remainder of the paper is as follows:

*Corresponding author
Received: 30 December 2012 / Accepted: 22 December 2013

Firstly, the new method is introduced for 2 and 3-term basis, then it is extended to the general case. Secondly, by presenting some numerical examples, the efficiency of both ordinary and chain least squares methods is compared.

2. Chain least squares method

2.1. Chain least squares with 2-term basis

Let $f \in L^2[a,b]$ and $\{L_i\}_{i=1}^2$ be a 2-term basis for a subspace of $L^2[a,b]$. One can approximate $f(s)$ by

$$f_2(s) = \sum_{i=1}^{2} a_i L_i(s), \quad s \in [a,b].$$

In least squares method for determining the unknown coefficients $\{a_i\}_{i=1}^2$, we solve the minimization problem

$$min_{a_1,a_2}\, e(a_1,a_2), \tag{3}$$

where

$$e(a_1,a_2) = \int_a^b [\sum_{i=1}^2 a_i L_i(s) - f(s)]^2 ds. \tag{4}$$

The ordinary approach for determining $\{a_i\}_{i=1}^2$, is to solve the normal equations (Datta, 2010).

$$\frac{\partial}{\partial a_i} e(a_1,a_2) = 0, \;\; i = 1,2. \tag{5}$$

The system (5) may be ill-conditioned in the sense that the condition number of the coefficient matrix be large and the coefficients $\{a_i\}_{i=1}^2$ be determined with large errors. For reducing the ill-conditioning of this problem, we proceed as follows.
Let $a = (a_1,a_2)^T$ be solution of the system (5). Then there exists constant $r \in \mathbb{R}$ such that

$$a_1 - a_2 = r.$$

It is clear that this equation characterizes a straight line trajectory in $\mathbb{R}^2$. We show this trajectory by g i.e.,

$$g(a_1,a_2) = a_1 - a_2 - r = 0. \tag{6}$$

It should be noted that the scalar r is unknown at this time. We suppose that the solution of (5) is settled on the trajectory (6). By this assumption, the minimization problem (3) is equivalent to

$$\begin{matrix} min & e(a_1,a_2) \\ s.t & g(a_1,a_2) = 0, \end{matrix}$$

then by the Lagrange multipliers method (Ito and Kunisch, 2008), we setup the problem

$$\begin{cases} \vec{\nabla} e = \lambda \vec{\nabla} g \\ g(a_1,a_2) = 0, \end{cases} \tag{7}$$

for $\lambda \in \mathbb{R}$, where

$$\vec{\nabla} e = \frac{\partial e}{\partial a_1}\vec{\imath} + \frac{\partial e}{\partial a_2}\vec{\jmath}, \;\; \vec{\nabla} g = \frac{\partial g}{\partial a_1}\vec{\imath} + \frac{\partial g}{\partial a_2}\vec{\jmath}.$$

By using (4) and (6), we obtain

$$\vec{\nabla} e = (2c_{11}a_1 + 2c_{12}a_2 - 2f_1)\vec{\imath} + (2c_{21}a_1 + 2c_{22}a_2 - 2f_2)\vec{\jmath},$$

and

$$\vec{\nabla} g = \vec{\imath} - \vec{\jmath},$$

where

$$c_{ij} = \int_a^b L_i(s)L_j(s)\,ds, \;\; f_i = \int_a^b L_i(s)f(s)ds, \quad i,j \in \{1,2\}.$$

Consequently, the system (7) is reduced to the system

$$\begin{cases} 2c_{11}a_1 + 2c_{12}a_2 - 2f_1 = \lambda \\ 2c_{21}a_1 + 2c_{22}a_2 - 2f_2 = -\lambda \\ a_1 - a_2 = r. \end{cases}$$

By summing the first two equations (to cancel λ), we obtain

$$\begin{cases} d_1a_1 + d_2a_2 = h \\ a_1 - a_2 = r, \end{cases} \tag{8}$$

where

$$h = \sum_{i=1}^{2} f_i, \;\; d_j = \sum_{i=1}^{2} c_{ij}, \;\; j = 1,2.$$

We solve the system (8), instead of (5) for determining the unknown coefficients $\{a_i\}_{i=1}^2$. From (8) the coefficients $\{a_i\}_{i=1}^2$ are determined in terms of the unknown scalar r as

$$a = DR, \tag{9}$$

where

$$a = (a_1,a_2)^T, \;\; R = (r,1)^T,$$

and

$$D = \frac{1}{N}\begin{pmatrix} d_2 & h \\ -d_1 & h \end{pmatrix}, \;\; N = d_1 + d_2.$$

Let $L(s) = (L_1(s), L_2(s))$. Then (4) can be written as

$$e(a_1,a_2) = \int_a^b [L(s)a - f(s)]^2 ds,$$

whence using (9), we obtain

$$E(r) = e(a_1, a_2) = \int_a^b [L(s)DR - f(s)]^2 ds.$$

Since

$$L(s)DR = L(s)D^1 r + L(s)D^2,$$

where

$$D^i = i'th\ column\ of\ D,$$

we have

$$E(r) = \int_a^b [L(s)D^1 r + L(s)D^2 - f(s)]^2 ds.$$

Let

$$p_i(s) = L(s)D^i,\ \ i = 1,2,$$

and

$$\bar{f}(s) = f(s) - p_2(s).$$

Then

$$E(r) = \int_a^b \left[p_1(s)r - \bar{f}(s)\right]^2 ds,$$

consequently, the minimization problem (3) is equivalent to

$$\min_{r\in\mathbb{R}} E(r).$$

For determining the scalar r, we put

$$\frac{dE}{dr} = E'(r) = 0,$$

then we get

$$r = \frac{\int_a^b p_1(s)\bar{f}(s)ds}{\int_a^b p_1^2(s)ds}.$$

Now by (9), the unknown coefficients $\{a_i\}_{i=1}^2$ are determined.

The following examples are given to show the efficiency of the modified method. Let f_2 be the least squares approximation of $f \in C[a,b]$ in the basis $\{L_i\}_{i=1}^2$. We also take

$$e(s) = |f_2(s) - f(s)|.$$

In the least squares approximation of the following examples, the basis functions are $\{1, s\}$.

Example 1. $f(s) = e^s,\ \ s \in [2, 2+\delta]$.

Example 2. $f(s) = sin(s),\ \ s \in [2, 2+\delta]$.

In the following tables, $\|e\|_\infty$ denote the infinity norm of error function i.e.,

$$\|e\|_\infty = max_{s\in[a,b]}|e(s)| \simeq max_{i=1,\dots,M}|e(s_i)|, \qquad (10)$$

where $s_i \in [a,b]$ are equidistant points. In numerical examples, we take $M = 11$. The numerical results show that the chain method operates better than the ordinary method in tiny domains (Tables 1 and 2).

Table 1. Maximum absolute errors, $\|e\|_\infty$, for Example 1

δ	ordinary method	chain method
10^{-01}	6.53×10^{-03}	6.53×10^{-03}
10^{-02}	6.19×10^{-05}	6.19×10^{-05}
10^{-03}	6.25×10^{-07}	6.16×10^{-07}
10^{-04}	6.34×10^{-09}	6.15×10^{-09}
10^{-05}	3.56×10^{-09}	6.15×10^{-11}
10^{-06}	2.73×10^{-08}	6.15×10^{-13}
10^{-07}	1.84×10^{-07}	6.21×10^{-15}
10^{-08}	2.44×10^{-08}	1.77×10^{-15}
10^{-09}	4.93×10^{-06}	8.88×10^{-16}
10^{-10}	2.92×10^{-05}	1.77×10^{-15}
10^{-11}	4.25×10^{-11}	1.77×10^{-15}
10^{-12}	3.51×10^{-03}	1.77×10^{-15}
10^{-13}	1.81×10^{-03}	1.77×10^{-15}
10^{-14}	3.26×10^{-01}	8.88×10^{-16}

Table 2. Maximum absolute errors, $\|e\|_\infty$, for Example 2

δ	Ordinary method	Chain method
10^{-01}	7.43×10^{-04}	7.43×10^{-04}
10^{-02}	7.56×10^{-06}	7.56×10^{-06}
10^{-03}	7.80×10^{-08}	7.57×10^{-08}
10^{-04}	7.69×10^{-10}	7.57×10^{-10}
10^{-05}	8.91×10^{-10}	7.57×10^{-12}
10^{-06}	1.31×10^{-09}	7.57×10^{-14}
10^{-07}	1.04×10^{-08}	5.55×10^{-16}
10^{-08}	3.07×10^{-09}	3.33×10^{-16}
10^{-09}	5.78×10^{-07}	1.11×10^{-16}
10^{-10}	2.15×10^{-06}	1.11×10^{-16}
10^{-11}	3.40×10^{-12}	2.22×10^{-16}
10^{-12}	1.69×10^{-04}	3.33×10^{-16}
10^{-13}	4.45×10^{-03}	1.11×10^{-16}
10^{-14}	3.23×10^{-02}	1.11×10^{-16}

Note: If well-behaved functions are approximated by least squares method in 2-term basis, the condition number of the normal equations will be small and the numerical solutions of this problem will be determined with high accuracy. In tiny domains, the condition number of these problems is large and the numerical solutions of the normal equations are determined with large errors. In this case, the efficiency of new method will appear. To show the efficiency of the new method in reducing

the condition number of normal equations, the condition numbers of both systems ((5) and (8)) for Example 1 are presented in Table 3.

Table 3. Condition numbers of Example 1

δ	ordinary method	chain method
10^{-01}	3.24×10^{04}	2.21×10^{00}
10^{-02}	3.02×10^{06}	2.21×10^{01}
10^{-03}	3.00×10^{08}	2.22×10^{02}
10^{-04}	3.00×10^{10}	2.22×10^{03}
10^{-05}	2.99×10^{12}	2.22×10^{04}
10^{-06}	2.99×10^{14}	2.22×10^{05}
10^{-07}	2.11×10^{16}	2.22×10^{06}
10^{-08}	∞	2.22×10^{07}
10^{-09}	1.35×10^{16}	2.22×10^{08}
10^{-10}	2.16×10^{16}	2.22×10^{09}
10^{-11}	1.73×10^{16}	2.22×10^{10}
10^{-12}	∞	2.22×10^{11}
10^{-13}	1.10×10^{16}	2.22×10^{12}
10^{-14}	∞	2.17×10^{13}

2.2. Chain least squares with 3*-term basis*

Let

$$f_3(s) = \sum_{i=1}^{3} a_i L_i(s), \quad s \in [a,b],$$

be the least squares approximation of $f \in L^2[a,b]$ on the basis of $\{L_i\}_{i=1}^3$. For achieving the unknown coefficients $\{a_i\}_{i=1}^3$, the following minimization problem must be solved

$$min_{a_1,a_2,a_3}\, e(a_1,a_2,a_3), \tag{11}$$

where

$$e(a_1,a_2,a_3) = \int_a^b [\sum_{i=1}^3 a_i L_i(s) - f(s)]^2 ds. \tag{12}$$

This minimization problem is equivalent to solving the following normal equations (Datta, 2010).

$$\frac{\partial}{\partial a_i} e(a_1,a_2,a_3) = 0, \;\; i = 1,2,3. \tag{13}$$

This problem may be ill-conditioned and the unknown coefficients $\{a_i\}_{i=1}^3$ are determined with large errors. For decreasing the ill-conditioning, we act as follows.

Let $a = (a_1,a_2,a_3)^T$ be the solution of the problem (13). Then there exist the real constants r_1 and r_2 such that

$$\begin{aligned} a_1 - a_2 &= r_1, \\ a_2 - a_3 &= r_2. \end{aligned}$$

It is clear that these equations characterize two planes in $\mathbb{R}^3$. We show these trajectories as follows.

$$\begin{aligned} g_1(a_1,a_2,a_3) &= a_1 - a_2 - r_1 = 0, \\ g_2(a_1,a_2,a_3) &= a_2 - a_3 - r_2 = 0, \end{aligned} \tag{14}$$

consequently, the minimization problem (11) can be written as

$$\begin{aligned} min \quad & e(a_1,a_2,a_3) \\ s.t \quad & g_1(a_1,a_2,a_3) = 0, \\ & g_2(a_1,a_2,a_3) = 0. \end{aligned}$$

By the lagrange multipliers method (Ito and Kunisch, 2008), we setup the problem

$$\begin{cases} \vec{\nabla} e = \lambda_1 \vec{\nabla} g_1 + \lambda_2 \vec{\nabla} g_2 \\ g_1(a_1,a_2,a_3) = 0 \\ g_2(a_1,a_2,a_3) = 0, \end{cases} \tag{15}$$

for $\lambda_1, \lambda_2 \in \mathbb{R}$. Since

$$\vec{\nabla} e = \frac{\partial e}{\partial a_1}\vec{\imath} + \frac{\partial e}{\partial a_2}\vec{\jmath} + \frac{\partial e}{\partial a_3}\vec{k},$$

using (12), we have

$$\begin{aligned} \vec{\nabla} e \;\; = \;\; & (2c_{11}a_1 + 2c_{12}a_2 + 2c_{13}a_3 - 2f_1)\vec{\imath} \\ + \;\; & (2c_{21}a_1 + 2c_{22}a_2 + 2c_{23}a_3 - 2f_2)\vec{\jmath} \\ + \;\; & (2c_{31}a_1 + 2c_{32}a_2 + 2c_{33}a_3 - 2f_3)\vec{k}, \end{aligned}$$

where

$$c_{ij} = \int_a^b L_i(s)L_j(s)\, ds, \;\; i,j \in \{1,2,3\},$$

and

$$f_i = \int_a^b L_i(s) f(s) ds, \;\; i \in \{1,2,3\}.$$

By (14) one gets

$$\vec{\nabla} g_1 = \vec{\imath} - \vec{\jmath}, \;\; \vec{\nabla} g_2 = \vec{\jmath} - \vec{k},$$

therefore, the system (15) is reduced to

$$\begin{cases} 2c_{11}a_1 + 2c_{12}a_2 + 2c_{13}a_3 - 2f_1 = \lambda_1 \\ 2c_{21}a_1 + 2c_{22}a_2 + 2c_{23}a_3 - 2f_2 = \lambda_2 - \lambda_1 \\ 2c_{31}a_1 + 2c_{32}a_2 + 2c_{33}a_3 - 2f_3 = -\lambda_2 \\ a_1 - a_2 = r_1 \\ a_2 - a_3 = r_2. \end{cases}$$

By summing the first three equations, we get (to get rid of scalars λ_1 and λ_2)

$$\begin{cases} d_1a_1 + d_2a_2 + d_3a_3 = h \\ a_1 - a_2 = r_1 \\ a_2 - a_3 = r_2, \end{cases} \tag{16}$$

where

$$h = \sum_{i=1}^{3} f_i, \quad d_j = \sum_{i=1}^{3} c_{ij}, \quad j \in \{1,2,3\}.$$

We solve the system (16) instead of (13) for determining the unknown coefficients $\{a_i\}_{i=1}^3$. From (16), the coefficients $\{a_i\}_{i=1}^3$ are determined in terms of the unknown scalars r_1 and r_2 as

$$a = DR, \tag{17}$$

where

$$a = (a_1, a_2, a_3)^T, \quad R = (r_1, r_2, 1)^T,$$

and

$$D = \frac{1}{N}\begin{pmatrix} d_1 + d_2 & d_3 & h \\ -d_1 & d_3 & h \\ -d_1 & -(d_1 + d_2) & h \end{pmatrix}, \quad N = \sum_{i=1}^{3} d_i.$$

Let

$$L(s) = (L_1(s), L_2(s), L_3(s)),$$

then from (12) we have

$$e(a_1, a_2, a_3) = \int_a^b [L(s)a - f(s)]^2 ds,$$

or equivalently (using (17))

$$E(r_1, r_2) = e(a_1, a_2, a_3) = \int_a^b [L(s)DR - f(s)]^2 ds.$$

Since

$$L(s)DR = L(s)D^1 r_1 + L(s)D^2 r_2 + L(s)D^3,$$

where

$$D^i = i'th\ column\ of\ D,$$

we have

$$E(r_1, r_2) = \int_a^b [L(s)D^1 r_1 + L(s)D^2 r_2 + L(s)D^3 - f(s)]^2 ds.$$

Let

$$p_i(s) = L(s)D^i, \quad i = 1,2,3,$$

and

$$\bar{f}(s) = f(s) - p_3(s),$$

then

$$E(r_1, r_2) = \int_a^b \left[r_1 p_1(s) + r_2 p_2(s) - \bar{f}(s)\right]^2 ds,$$

consequently, the minimization problem (11) is equivalent to

$$min_{r_1, r_2}\, E(r_1, r_2). \tag{18}$$

For determining the unknown scalars r_1 and r_2, we must solve the least squares problem (18). If this problem is ill-posed we solve it by the chain method. Otherwise, it is solved by the ordinary method. After determining the scalars r_1 and r_2, the unknown coefficients $\{a_i\}_{i=1}^3$ will be obtained from (17).

2.3. Chain least squares with n-term basis

Let f_n be the least squares approximation of $f \in L^2[a, b]$ on the basis $\{L_i\}_{i=1}^n$, i.e.,

$$f_n(s) = \sum_{i=1}^{n} a_i L_i(s), \quad s \in [a, b].$$

To determine the unknown coefficients $\{a_i\}_{i=1}^n$ we must solve the following minimization problem

$$min_{a_1, \dots, a_n}\, e(a_1, \dots, a_n), \tag{19}$$

where

$$e(a_1, \dots, a_n) = \int_a^b [\sum_{i=1}^n a_i L_i(s) - f(s)]^2 ds. \tag{20}$$

This is equivalent to solving the following normal equations (Datta, 2010).

$$\frac{\partial}{\partial a_i} e(a_1, \dots, a_n) = 0, \quad i = 1, \dots, n. \tag{21}$$

This system is often an ill-conditioned problem to solve. For reducing the ill-conditioning, we act as follows.

Let $a = (a_1, \dots, a_n)^T$ be the solution of the problem (21). Then there exist scalars $\{r_i\}_{i=1}^{n-1}$ belonging to $\mathbb{R}$ such that

$$a_i - a_{i+1} = r_i, \quad i = 1, \dots, n-1.$$

In other words, the solution of the system (21) is the intersection of the following $n-1$ surfaces:

$$g_i(a_1, \dots, a_n) = 0, \quad i = 1, \dots, n-1,$$

where

$$g_i(a_1, \dots, a_n) = a_i - a_{i+1} - r_i, i = 1, \dots, n-1, \tag{22}$$

therefore, the minimization problem (19) is equivalent to

$$\begin{aligned} min \quad & e(a_1, \dots, a_n) \\ s.t \quad & g_i(a_1, \dots, a_n) = 0, \quad i = 1, \dots, n-1. \end{aligned} \tag{23}$$

By the lagrange multipliers method (Ito and Kunisch, 2008), there exist scalars $\{\lambda_i\}_{i=1}^{n-1}$ belonging to $\mathbb{R}$ such that the problem (23) is equivalent to

$$\begin{cases} \vec{\nabla} e = \sum_{i=1}^{n-1} \lambda_i \vec{\nabla} g_i, \\ g_i(a_1, \dots, a_n) = 0, \ \ i = 1, \dots, n-1. \end{cases} \tag{24}$$

Since

$$\vec{\nabla} e = \left(\frac{\partial e}{\partial a_1}, \dots, \frac{\partial e}{\partial a_n} \right),$$

by (20) and (22) we have

$$\frac{\partial e}{\partial a_i} = 2 \sum_{j=1}^{n} c_{ij} - 2 f_i, \ \ i = 1, \dots, n,$$

where

$$c_{ij} = \int_a^b L_i(s) L_j(s)\, ds, \ \ i,j \in \{1, \dots, n\},$$

$$f_i = \int_a^b L_i(s) f(s) ds\,, \ \ i \in \{1, \dots, n\},$$

and

$$\vec{\nabla} g_i = (0, \dots, 0, 1, -1, 0, \dots, 0), \ \ i = 1, \dots, n-1,$$

where 1 and -1 are the i'th and $(i+1)$'th components of $\vec{\nabla} g_i$ respectively. Consequently, the system (24) is equivalent to

$$\begin{cases} 2c_{11}a_1 + 2c_{12}a_2 + \cdots + 2c_{1n}a_n - 2f_1 = \ \ \lambda_1 \\ 2c_{21}a_1 + 2c_{22}a_2 + \cdots + 2c_{2n}a_n - 2f_2 = \lambda_2 - \lambda_1 \\ \quad \vdots \qquad\qquad \cdots \qquad\qquad \vdots \\ 2c_{n1}a_1 + 2c_{n2}a_2 + \cdots + 2c_{nn}a_n - 2f_n = -\lambda_{n-1} \\ a_1 - a_2 = r_1 \\ \qquad\qquad \ddots \qquad\qquad \vdots \\ a_{n-1} - a_n = r_{n-1}. \end{cases}$$

By summing the first n equations, we have (for removing scalars $\lambda_1, \dots, \lambda_{n-1}$)

$$\begin{cases} d_1 a_1 + d_2 a_2 + \cdots + d_n a_n = h \\ a_1 - a_2 = r_1 \\ \qquad\qquad \cdots \\ a_{n-1} - a_n = r_{n-1} \end{cases} \tag{25}$$

where

$$h = \sum_{i=1}^{n} f_i, \ \ d_j = \sum_{i=1}^{n} c_{ij}\,, \ \ j \in \{1, \dots, n\}.$$

We solve the system (25) instead of (21) for determining the unknown coefficients $\{a_i\}_{i=1}^{n}$. Because (25) is often more well conditioned than (21). From the system (25) the coefficients $\{a_i\}_{i=1}^{n}$ are determined in terms of the unknown scalars $\{r_i\}_{i=1}^{n-1}$ as follows:

$$a = DR, \tag{26}$$

where

$$a = (a_1, \dots, a_n)^T, \ \ R = (r_1, \dots, r_{n-1}, 1)^T,$$

and

$$D = \frac{1}{N} \begin{pmatrix} N - t_1 & N - t_2 & \cdots & N - t_{n-1} & h \\ -t_1 & N - t_2 & \cdots & N - t_{n-1} & h \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ -t_1 & -t_2 & \cdots & N - t_{n-1} & h \\ -t_1 & -t_2 & \cdots & -t_{n-1} & h \end{pmatrix},$$

where

$$N = \sum_{i=1}^{n} d_i,$$

$$t_1 = d_1, \ \ t_i = t_{i-1} + d_i, \ \ i = 1, \dots, n-1.$$

Let

$$L(s) = (L_1(s), \dots, L_n(s)),$$

then from (20) we have

$$e(a_1, \dots, a_n) = \int_a^b [L(s)a - f(s)]^2 ds,$$

or equivalently (using (26))

$$\begin{aligned} E(r_1, \dots, r_{n-1}) &= e(a_1, \dots, a_n) \\ &= \int_a^b [L(s)DR - f(s)]^2 ds, \end{aligned}$$

Since

$$\begin{aligned} L(s)DR = L(s)D^1 r_1 + \cdots + L(s)D^{n-1} r_{n-1} \\ + L(s)D^n, \end{aligned}$$

where

$$D^i = i'th\ column\ of\ D,$$

one can write

$$\begin{aligned} E(r_1, \dots, r_{n-1}) = \int_a^b [L(s)D^1 r_1 + \cdots \\ + L(s)D^{n-1} r_{n-1} + L(s)D^n \\ - f(s)]^2 ds. \end{aligned}$$

Let

$$p_i(s) = L(s)D^i, \ \ i = 1, \dots, n,$$

and

$$\bar{f}(s) = f(s) - p_n(s),$$

then

$$E(r_1, \dots, r_{n-1}) = \int_a^b \left[r_1 p_1(s) + \cdots + r_{n-1} p_{n-1}(s) - \bar{f}(s)\right]^2 ds,$$

and so the minimization problem (19) is equivalent to

$$min_{r_1,\dots,r_{n-1}} E(r_1, \dots, r_{n-1}). \quad (27)$$

For determining the unknown scalars $\{r_i\}_{i=1}^{n-1}$, we must solve the $(n-1)$-term least squares problem (27). If it's normal equations are well-conditioned, it is solved by ordinary least squares method, otherwise, it is reduced to an $(n-2)$-term least squares problem. This trend is continued so that we reach a well-conditioned problem. We define this approach as chain least squares method. In all the presented numerical examples we have continued the chain method up to the final stage in order to get the maximum accuracy without doing any comparison.

The following numerical results show that, the chain method often operates better than the ordinary method. In the least squares approximation of the following examples, we take $\{1, s, \dots, s^n\}$ as basis functions. The maximum of absolute errors is computed as (10).

Example 3. $f(s) = cos\left(\frac{1}{4}s\right)$, $s \in \left[0, \frac{\pi}{2}\right]$.

Example 4. $f(s) = ln(s)$, $s \in \left[1, \frac{3}{2}\right]$.

Example 5. $f(s) = sinh(s)$, $s \in [0,1]$.

The maximum absolute errors of Examples 3, 4 and 5 by the both methods are reported in Tables 4, 5 and 6 respectively.

Table 4. Maximum absolute errors, $\|e\|_\infty$, for Example 3

n	Ordinary method	Chain method
1	1.26×10^{-02}	1.26×10^{-02}
2	1.12×10^{-04}	1.12×10^{-04}
3	1.39×10^{-05}	1.39×10^{-05}
4	6.55×10^{-08}	6.55×10^{-08}
5	5.41×10^{-09}	5.41×10^{-09}
6	1.79×10^{-11}	1.72×10^{-11}
7	3.30×10^{-12}	1.07×10^{-12}
8	1.95×10^{-11}	2.99×10^{-15}
9	2.25×10^{-10}	4.44×10^{-16}

Table 5. Maximum absolute errors, $\|e\|_\infty$, for Example 4

n	Ordinary method	Chain method
1	1.47×10^{-02}	1.47×10^{-02}
2	1.20×10^{-03}	1.20×10^{-03}
3	1.04×10^{-04}	1.04×10^{-04}
4	9.37×10^{-06}	9.37×10^{-06}
5	8.62×10^{-07}	8.62×10^{-07}
6	7.79×10^{-08}	8.05×10^{-08}
7	1.37×10^{-07}	7.59×10^{-09}
8	1.42×10^{-07}	7.23×10^{-10}
9	1.00×10^{-07}	7.53×10^{-11}

Table 6. Maximum absolute errors, $\|e\|_\infty$, for Example 5

n	Ordinary method	Chain method
1	5.40×10^{-02}	5.40×10^{-02}
2	9.88×10^{-03}	9.88×10^{-03}
3	3.52×10^{-04}	3.52×10^{-04}
4	3.85×10^{-05}	3.85×10^{-05}
5	3.85×10^{-05}	3.85×10^{-05}
6	6.67×10^{-08}	6.67×10^{-08}
7	1.07×10^{-09}	1.07×10^{-09}
8	7.02×10^{-11}	6.51×10^{-11}
9	4.06×10^{-11}	4.44×10^{-13}

3. Numerical solution of some functional equations

In this section, we consider some examples of the ill-posed functional equations of the form (1) in order to compare the accuracy of both methods. Also we take $\{L_i\}_{i=1}^n = \{t^{i-1}\}_{i=1}^n$.

Case 1.

$$(Ax)(s) = \int_a^b k(s,t)x(t)dt, \quad s \in [a,b].$$

In this case, we deal with Fredholm integral equations of the first kind (Bitsadze, 1995; Delves and Mohamed, 1985). Many engineering problems cane be modeled by this kinds of equations. For example, one and two-dimensional scattering from conducting bodies can be modeled by them (Balanis, 1989). In most numerical approaches for solving these equations, the attempt is to get accurate approximations by reducing their ill-conditioning (Babolian and Delves, 1979; Babolian et al., 2007; Groetsch, 1984; Maleknejad et al., 2006).

Case 2.

$$(Ax)(s) = \varepsilon x(s) - \int_a^b k(s,t)x(t)dt, \quad s \in [a,b].$$

In this case, we face the singularly perturbed linear Fredholm integral equations of the second

kind (Smith, 1985; Sweilam et al., 2009). These ill-posed equations ($\mathcal{E} \to 0$) arise in various fields of science and engineering (Sweilam et al., 2009).

Note: For computing the related integrals, we use a Gaussian quadrature rule of order 10 with 32 significant digits.
The numerical examples are as follows:

Example 6.

$$\int_0^1 e^{st}x(t)dt = \frac{e^{s+1}-1}{s+1}, \quad s \in [0,1].$$

Example 7.

$$\int_0^2 sin(st)\, x(t)dt = 2\frac{(4s+1)\, sin\left(2s-\frac{1}{2}\right) - (4s-1)sin(2s+\frac{1}{2})}{16s^2-1},$$
$s \in [0,2]$.

Example 8.

$$\int_0^3 e^{tsin(s)}x(t)dt = \frac{(3\, sin(s)-1)e^{3\, sin(s)}+1}{sin^2(s)},$$
$s \in [0,3]$.

The exact solutions are e^t, $sin(\frac{1}{4}t)$ and t respectively. The maximum absolute errors of these examples using both methods are reported in Tables 7, 8 and 9.

Example 9.

$$\varepsilon x(s) - \int_{-1}^1 cosh(s+t)\, x(t)dt = -cosh(s)\,,$$
$s \in [-1,1]$,

which has the exact solution

$$x(t) = \frac{2cosh(t)}{2+sinh(2)-2\varepsilon}.$$

We solve this perturbed problem for the values of $\varepsilon = 10^{-3}$, 10^{-4} and 10^{-5}. The errors of dynamical systems method (DSM), proposed in (Sweilam et al., 2009), are about 2.9×10^{-3} for these values of ε, while the error of our presented method is 5.6×10^{-10} for $n = 9$ (Tables 10, 11 and 12).

Table 7. Maximum absolute errors, $\|e\|_\infty$, for Examples 6

n	Ordinary method	Chain method
1	9.16×10^{-01}	9.16×10^{-01}
2	1.42×10^{-01}	1.42×10^{-01}
3	1.46×10^{-02}	1.46×10^{-02}
4	1.05×10^{-03}	1.05×10^{-03}
5	1.23×10^{-02}	5.94×10^{-05}
6	3.77×10^{-01}	2.71×10^{-06}
7	7.53×10^{-02}	2.52×10^{-06}

Table 8. Maximum absolute errors, $\|e\|_\infty$, for Examples 7

n	Ordinary method	Chain method
1	2.76×10^{-01}	2.76×10^{-01}
2	8.14×10^{-03}	8.14×10^{-03}
3	2.83×10^{-03}	2.83×10^{-03}
4	4.07×10^{-05}	4.07×10^{-05}
5	4.73×10^{-04}	6.77×10^{-06}
6	1.45×10^{-02}	6.45×10^{-08}
7	2.61×10^{-03}	1.83×10^{-08}

Table 9. Maximum absolute errors, $\|e\|_\infty$, for Example 8

n	Ordinary method	Chain method
1	2.08×10^{-00}	2.08×10^{-00}
2	1.12×10^{-13}	1.12×10^{-16}
3	1.39×10^{-11}	1.39×10^{-13}
4	6.55×10^{-07}	6.55×10^{-12}
5	5.41×10^{-04}	5.41×10^{-10}
6	1.79×10^{-03}	1.72×10^{-09}
7	3.30×10^{-02}	1.07×10^{-08}

Table 10. Maximum absolute errors, $\|e\|_\infty$, for Example 9 with $\varepsilon = 10^{-3}$

n	Ordinary method	Chain method
1	1.23×10^{-01}	1.23×10^{-01}
2	1.23×10^{-01}	1.23×10^{-01}
3	3.57×10^{-03}	3.57×10^{-03}
4	3.57×10^{-03}	3.57×10^{-03}
5	3.55×10^{-05}	3.55×10^{-05}
6	3.55×10^{-05}	3.55×10^{-05}
7	1.81×10^{-07}	1.80×10^{-07}
8	1.83×10^{-07}	1.80×10^{-07}
9	6.03×10^{-08}	5.57×10^{-10}
10	6.92×10^{-08}	5.57×10^{-10}

Table 11. Maximum absolute errors, $\|e\|_\infty$, for Example 9 with $\varepsilon = 10^{-4}$

n	Ordinary method	Chain method
1	1.23×10^{-01}	1.23×10^{-01}
2	1.23×10^{-01}	1.23×10^{-01}
3	3.57×10^{-03}	3.57×10^{-03}
4	3.57×10^{-03}	3.57×10^{-03}
5	3.55×10^{-05}	3.55×10^{-05}
6	3.55×10^{-05}	3.55×10^{-05}
7	2.24×10^{-06}	1.80×10^{-07}
8	2.68×10^{-06}	1.80×10^{-07}
9	9.53×10^{-06}	5.60×10^{-10}
10	1.20×10^{-05}	5.60×10^{-10}

Table 12. Maximum absolute errors, $\|e\|_\infty$, for Example 9 with $\varepsilon = 10^{-5}$

n	Ordinary method	Chain method
1	1.23×10^{-01}	1.23×10^{-01}
2	1.23×10^{-01}	1.23×10^{-01}
3	3.56×10^{-03}	3.57×10^{-03}
4	3.56×10^{-03}	3.57×10^{-03}
5	2.49×10^{-05}	3.55×10^{-05}
6	3.86×10^{-05}	3.55×10^{-05}
7	3.09×10^{-04}	1.80×10^{-07}
8	3.54×10^{-04}	1.80×10^{-07}
9	2.09×10^{-03}	5.65×10^{-10}
10	2.11×10^{-03}	5.49×10^{-10}

At the end of this section, by making random noise data on the right hand side of the examples 3.1 and 3.3 and solving these equations by the chain least squares method, the stability of new method is investigated. Numerical results are shown in the Tables 13 and 14. It should be mentioned that, for perturbations of order $\xi = 10^{-m}$ in data of a problem, the maximum accuracy of the approximate solutions will be of order $\xi = 10^{-m}$ (for a positive integer m).

Table 13. Maximum absolute errors, $\|e\|_\infty$, of Example 6 for random noise data by new method

$\xi = 10^{-5}$	$\xi = 10^{-6}$	$\xi = 10^{-7}$
1	9.16×10^{-01}	9.16×10^{-01}
2	1.41×10^{-01}	1.42×10^{-01}
3	2.28×10^{-02}	1.54×10^{-02}
4	1.39×10^{-02}	1.25×10^{-03}
5	2.38×10^{-02}	2.43×10^{-03}
6	3.44×10^{-02}	3.44×10^{-03}
7	4.72×10^{-02}	4.72×10^{-03}
$\xi = 10^{-5}$	$\xi = 10^{-6}$	$\xi = 10^{-7}$
9.16×10^{-01}	9.16×10^{-01}	9.16×10^{-01}
1.42×10^{-01}	1.42×10^{-01}	1.42×10^{-01}
1.47×10^{-02}	1.46×10^{-02}	1.46×10^{-02}
9.72×10^{-04}	1.04×10^{-03}	1.05×10^{-03}
2.96×10^{-04}	8.31×10^{-05}	6.17×10^{-05}
3.42×10^{-04}	3.17×10^{-05}	2.77×10^{-06}
4.71×10^{-04}	4.71×10^{-05}	7.95×10^{-06}

Table 14. Maximum absolute errors, $\|e\|_\infty$, of Example 8 for random noise data by new method

n	$\xi = 10^{-3}$	$\xi = 10^{-4}$
1	2.08×10^{-00}	2.08×10^{-00}
2	7.74×10^{-04}	7.74×10^{-05}
3	2.18×10^{-03}	2.18×10^{-04}
4	4.27×10^{-03}	4.27×10^{-04}
5	7.05×10^{-03}	7.05×10^{-04}
6	1.04×10^{-02}	1.04×10^{-03}
7	1.44×10^{-02}	1.44×10^{-03}
$\xi = 10^{-5}$	$\xi = 10^{-6}$	$\xi = 10^{-7}$
2.08×10^{-00}	2.08×10^{-00}	2.08×10^{-00}
7.74×10^{-06}	7.74×10^{-07}	7.74×10^{-08}
2.18×10^{-05}	2.18×10^{-06}	2.18×10^{-07}
4.27×10^{-05}	4.27×10^{-06}	4.27×10^{-07}
7.05×10^{-05}	7.05×10^{-06}	7.05×10^{-07}
1.04×10^{-04}	1.04×10^{-05}	1.04×10^{-06}
1.44×10^{-04}	1.44×10^{-05}	1.40×10^{-06}

4. Conclusion

The numerical results show that the chain least squares method is often more accurate and stable than the ordinary least squares approach. Reduction of the n-term least squares problem in the $(n-1)$-term problem is a new approach that can be the topic of new researches. It should be mentioned that the defined surfaces (22) have a main role in increasing the efficiency of least squares approximations. For decreasing the ill-conditioning of the least squares problem better than the presented method, finding other trajectories is suggested as an open problem.

References

Aksan, E. N., Ozdes, A., & Ozis, T. (2006). A numerical solution of Burgers equation based on least squares approximation. *Applied Mathematics and Computation*, 176, 270–279.

Alexander, P., & George, F. P. (1990). A parallel least squares collocation conjugate gradient approach for the advection diffusion equation. *Advances in Water Resource*s, 13, 126–136.

Babolian, E., & Delves, L. M. (1979). An Augmented Galerkin Method for First Kind Fredholm Equations. *J. Inst. Maths. Applics*., 24, 157–174.

Babolian, E., Lotfi, T., & Paripour, M. (2007). Wavelet moment method for solving Fredholm integral equations of the first kind. *Applied Mathematics and Computation*, 186, 1467–1471.

Balanis, C. A. (1989). *Advanced Engineering Electromagnetics*. Wiley, New York.

Bitsadze, A. V. (1995). *Integral Equations of First Kind*. World Scientific Publishing Co. Pte. Ltd.

Ching, L. C., & Suh-Yuh Y. (2002). Analysis of the L^2 least squares finite element method for the velocity vorticity pressure Stokes equations with velocity boundary conditions. *Applied Mathematics and Computation*, 130, 121–144.

Datta, B. N. (2010). Numerical Linear Algebra and Applications. Second Edition, SIAM.

Delves, L. M., & Mohamed, J. L. (1985). *Computational Methods for Integral Equations*. Cambridge University Press.

Groetsch, C. W. (1984). *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*. Research Notes in Mathematics, vol. 105, Pitman, Boston.

Ito, K., & Kunisch, K. (2008). Lagrange multipliers approach to variational problems and applications. *SIAM*.

Jagalur-Mohan, J., Feijo, G., & Oberai, A. (2013). A Galerkin, least squares method for time harmonic Maxwell equations using Nédélec elements. *Journal of Computational Physics*, 235, 67–81.

Jannike, S., & Hugo, A. J. (2012). Effects of Jacobi polynomials on the numerical solution of the pellet equation using the orthogonal collocation, Galerkin, tau and least squares methods. *Computers & Chemical Engineering*, 39, 1–21.

Jinming, W. (2012). Least squares methods for solving partial differential equations by using Bzier control points. *Applied Mathematics and Computation*, 219, 3655–3663.

Kincaid, D. R., & Ward, C. E. (2002). *Numerical Analysis: Mathematics of scientific computing*. American Mathematical Society.

King, B. B., & Krueger, D. A. (2003). Burgers' equation: Galerkin least-squares approximations and feedback control. *Mathematical and Computer Modelling*, 38, 1075–1085.

Laeli, D. H., & Maalek, G. F. M. (2012). Numerical solution of Volterra − Fredholm integral equations by moving least square method and Chebyshev polynomials. *Applied Mathematical Modelling*, 36, 3283–3288.

Maleknejad, K., Aghazadeh, N., & Mollapourasl, R. (2006). Numerical solution of Fredholm integral equation of the first kind with collocation method and estimation of error bound. *Applied Mathematics and Computation*, 179, 352–359.

Nashed, M. N. (1976). On Moment-Discretization and Least-Squares Solutions of Linear Integral Equations of the First Kind. *J. Math. Anal. Appl*., 53, 359–366.

Smith, O. R. (1985). *Singularly Perturbed Theory*. Cambridge University Press, New York.

Sweilam, N. H., Nagy, A. M., & Alnasr, M. H. (2009). An efficient dynamical systems method for solving singularly perturbed integral equations with noise. *Computers & Mathematics with Applications*, 58, 1418–1424.