Journal of Computational Mathematics, Vol.10, No.1, 1992, 29–38.

SOLUTION OF AN OVERDETERMINED SYSTEM OF LINEAR EQUATIONS IN L_2, L_∞, L_p NORM USING L.S. TECHNIQUES *

Yang Shu-guang Liao Jian-wen
(*Wuhan University, Wuhan, Hubei, China*)

Abstract

A lot of curve fitting problems of experiment data lead to solution of an overdetermined system of linear equations. But it is not clear prior to that whether the data are exact or contaminated with errors of an unknown nature. Consequently we need to use not only L_2-solution of the system but also L_∞-or L_p-solution.

In this paper, we propose a universal algorithm called the Directional Perturbation Least Squares (DPLS) Algorithm, which can give optimal solutions of an overdetermined system of linear equations in L_2, L_∞, L_p $(1 \le p < 2)$ norms using only L.S. techniques (in §2). Theoretical principle of the algorithm is given in §3. Two examples are given in the end.

§1. Introduction

Let us assume that

$$X_i = (x_{i_1}, \cdots, x_{i_k})^T \in R^k, \quad Y_i = f(X_i), \quad i = 1, 2, \cdots, m$$

are the given original data. We wish to find a fitting formula linearly depending on some parameters $b = (b_1, \cdots, b_n)^T$

$$F(b, X) = \sum_{j=1}^{n} b_j \phi_j(X) = [\Phi(X)]^T b \tag{1.1}$$

such that

$$\|r(b)\|_p = \|Y - \tilde{F}(b, X)\|_p = \min_{b \in R^n} \tag{1.2}$$

* Received September 7, 1988.

where

$$Y = (Y_1, \cdots, Y_m)^T \in R^m, \quad b = (b_1, \cdots, b_n)^T \in R^n,$$
$$\tilde{F}(b, X) = (F(b, X_1), \cdots, F(b, X_m))^T \in R^m,$$
$$r(b) = Y - \tilde{F}(b, X),$$

$\left\{\phi_i(X)\right\}_1^n$: n linearly independent given functions on the discrete set $\{X_i\}_{i=1}^m$,

$\|\cdot\|_p$: p-norm of a vector $\|r\|_p = (\sum_{i=1}^{m} |r_i|^p)^{1/p}$, $1 \leq p \leq \infty$.

Problem (1.2) can be converted to the equivalent problem of solving an overdetermined linear system

$$AX = Y \tag{1.3}$$

by minimizing the p-norm (commonly p=1,2,∞) of the residual $r(X) = Y - AX$, i.e, to find a vector $b \in R^n$ such that

$$\|Y - Ab\|_p = \min_{X \in R^n} \|Y - AX\|_p \tag{1.4}$$

where

$$A = \begin{bmatrix} \phi_1(X_1) & \cdots & \phi_n(X_1) \\ \cdots & \cdots & \cdots \\ \phi_1(X_m) & \cdots & \phi_n(X_m) \end{bmatrix} \in R^{m \times n}.$$

Different norms normally lead to different solution b . If $p = 2$, it is well-known that

$$b = A^+ Y$$

where A^+ is the pseudo-inverse of A obtained by any method, for instance the S.V.D. method. When rank $(A) \equiv r_A = n$, b is the unique solution, and when $r_A < n$, b is a uniquely definite minimal 2-norm solution; but if $p = 1$ or $p = \infty$, the problems are complicated because the function $f(X) = \|Y - AX\|_p$ is not differentiable for those values of p. Indeed , there are several good techniques available for 1-norm and ∞-norm minimization [2]. But those techniques are more complex and very different from the L.S. Technique. As J.R. Rice ([1] p.371) indicates, "the theories behind both L_1 and Chebyshev (L_∞) approximation are too complex and difficult to present here. The numerical methods that have been developed from these theories are more complex and require more computation than the methods for least squares, but one could hardly hope for them to be as simple and easy as for least squares. On the other hand, these methods are reasonably efficient, and it is practical to use them in a wide variety of applications. If the problem at hand is such that least squares might not be entirely appropriate (e.g, one may really want to minimize the maximum error or the data might be contaminated significantly with errors of an unknown nature), then one should try one of the programs discussed later and see how effective these other two norms are." These are the motivations which lead to the proposed universal algorithm using only L.S. techniques. Most of the methods provided in the literature are divided in two types. One is the ascent method. The earliest was due to Stiefel (1959), called the exchange method, in which a single index is exchanged at each iteration to find an extremal subset. Since there is only a finite number of such subsets, the type is a finite method; based on the linear programming approach. The other is the descent method; there are a version due to Cline (1976) and a version improved by Bartels, Conn, and Charalambous (1978). We describe simply the latter method.

If a is an initial approximate solution of (1.2), we wish to find a corrective vector $c \in R^n$ and a small number $\gamma \in R$ such that $\|r(a+\gamma c)\|_\infty < \|r(a)\|_\infty$. The existence of vector c and number γ is guaranteed by Theorem 2.7 of [2] (p.44).

The directional perturbation method proposed in this paper is also a descent method, but we always take initial approximate solution $b^{(0)} = A^+Y = A^+Y^{(0)}$. Using linear dependence of b with respect to Y, we add some suitable perturbation vectors $\delta_k Y$ to Y such that $\|r(A^+(Y+\delta_k Y))\|_\infty < \|r(A^+Y)\|_\infty$ and $b^{(k)} = A^+Y^{(k)} = A^+(Y+\delta_k Y)$ tends to L_∞-norm optimal solution (or T-solution) of problem (1.2) when $k = 1, 2, \cdots, k_0 (\le n)$.

On the other hand, we shall give some suitable perturbation matrices $\Sigma^{(k-1)}$ depending on the residual $r^{(k-1)}$ such that

$$X^{(k)} = (A^{(k)})^+ Y^{(k)} \quad \text{(L.S-solution)}$$

where

$$A^{(k)} = \Sigma^{(k-1)} A \ , \ Y^{(k)} = \Sigma^{(k-1)} Y \tag{1.5}$$

which tends to an L_p-norm optimal solution of the system (1.3) (i.e, of problem (1.2)) $(k = 1, 2, \cdots)$.

§2. D.P.L.S-Algorithm

Given a system

$$AX = Y, \quad A \in R_r^{m\times n}, \quad X \in R^n, \quad Y \in R^m \tag{2.1}$$

where

$$r = \operatorname{rank}(A) \le n \le m.$$

1° Calculate A^+ and

$$X^{(0)} = A^+Y = A^+Y^{(0)}. \tag{2.2}$$

2° Calculate

$$r^{(k-1)} = Y - AX^{(k-1)} \text{ and } \|r^{(k-1)}\|_p, \quad p = 2,\ \infty,\ 1. \tag{2.3}$$

If $\|r^{(k-1)}\|_\infty < \varepsilon$, then stop; else go to 3°.

3° Determine the set

$$M^{(k-1)} = \{i : |r_i^{(k-1)}| = d^{(k-1)}, i \in E\} = \{i_1, i_2, \cdots, i_{p_{k-1}}\} \tag{2.4}$$

where $d^{(k-1)} = \max\limits_{i\in E} |r_i^{(k-1)}|,\ \ E = \{1, 2, \cdots, m\},\ \ 1 \le p_{k-1} \le m.$

4° Calculate the perturbation vector

$$a^{(k)} = (a_1^{(k)}, \cdots, a_{p_{k-1}}^{(k)})^T = \rho\alpha^{(k)} = \rho(\alpha_1^{(k)}, \cdots, \alpha_{p_{k-1}}^{(k)})^T$$

where $\alpha^{(k)}$ is the solution of the following system of order p_{k-1}:

$$\Delta\alpha^{(k)} = \bar{r}^{(k-1)} \tag{2.5}$$

where

$$\Delta = \begin{bmatrix} \delta_{i_1 i_1} & \cdots & \delta_{i_1 i_{p_{k-1}}} \\ \cdots & \cdots & \cdots \\ \delta_{i_{p_{k-1}} i_1} & \cdots & \delta_{i_{p_{k-1}} i_{p_{k-1}}} \end{bmatrix}, \tag{2.6}$$

$$\delta_{i_q i_h} = e_{i_q}^T AA^+ e_{i_h} \quad (q, h = 1, \cdots, p_{k-1}),$$

$$\bar{r}^{(k-1)} = (r_{i_1}^{(k-1)},\ \cdots,\ r_{i_{p_{k-1}}}^{(k-1)})^T$$

where e_i is the i-th basis of R^m.

The perturbation parameter ρ is determined by the following condition:

$$\max_{i\in E-M^{(k-1)}} |r_i^{(k-1)} - \rho \sum_{h=1}^{p_{k-1}} \alpha_h^{(k)} \delta_{ii_h}| = (1-\rho)d^{(k-1)}, \quad 0 < \rho < 1 \tag{2.7}$$

where

$$\delta_{ii_h} = e_i^T AA^+ e_{i_h},\ i \in E - M^{(k-1)},\ i_h \in M^{(k-1)}.$$

5° Add the perturbation vector to Y

$$Y^{(k)} = Y^{(k-1)} + \sum_{h=1}^{p_{k-1}} a_h^{(k)} e_{i_h} \tag{2.8}$$

and calculate the perturbation solution

$$X^{(k)} = A^{+}Y^{(k)} = X^{(k-1)} + \sum_{h=1}^{p_{k-1}} a_h^{(k)} A^{+} e_{i_h}. \tag{2.9}$$

6° If $p_k > n$ (or rank(A)) or $k > n$ (or rank(A)), then $X^{(k)}$ is the L_∞-solution of the system (2.1), go to step 7°; if $p_k < n$ (or rank(A)) or $k < n$ (or rank(A)) and det(Δ)=0, it means that $X^{(k)}$ is a degenerate L_∞-solution, then go to step 7°; else $k := k + 1$, go to step 2°.

7° In order to obtain an (approximately) optimal L_p-solution $(1 \le p < 2)$ of the system (2.1), make use of the residual vector $r^{(0)}$ of the system (2.1) with respect to L_2-solution $X^{(0)}$ $(r^{(0)} = Y - AX^{(0)})$ to get

$$\sigma_i^{(k-1)} = \begin{cases} \dfrac{1}{|r_i^{(k-1)}|^{2-p}}, & \text{if} |r_i^{(k-1)}|^{2-p} > \varepsilon, \\ \dfrac{1}{\varepsilon}, & \text{if} |r_i^{(k-1)}|^{2-p} \le \varepsilon \end{cases} \tag{2.10}$$

and form the perturbation matrix

$$\Sigma^{(k-1)} = \begin{bmatrix} (\sigma_1^{(k-1)})^{1/2} & & 0 \\ & \ddots & \\ 0 & & (\sigma_m^{(k-1)})^{1/2} \end{bmatrix}. \tag{2.11}$$

8° Multiply both members of the system (2.1) by $\Sigma^{(k-1)}$

$$\Sigma^{(k-1)}AX = \Sigma^{(k-1)}Y$$

and denote it by

$$A^{(k)}X = Y^{(k)}.$$

9° Calculate $X^{(k)} = (A^{(k)})^{+}Y^{(k)}$ and $\|r^{(k)}\|_p$.

10° If $\|X^{(k)} - X^{(k-1)}\|_\infty < \varepsilon$, then we shall take $X^{(k)}$ as an (approximately) optimal L_p-solution and stop; else $k := k + 1$ and go to 7°.

§3. Theoretical Results

The steps 1° − 2° of the algorithm are shared for L_2, L_∞, and L_p norm solution, but steps 3° − 6° serve only for L_∞-solution, and in this case we have

Lemma 1. *If the system of vectors* $\{\Phi(X_{i_h})\}$ $(h = 1, \cdots, p_{k-1})$ *is linearly independent, then the system of linear equations* (2.5) *has one and only one solution* $\alpha^{(k)}$.

Proof. It can be proved that the matrix Δ is symmetric positive definite (see [4]).

Theorem 2. *For $X^{(k-1)}$ which satisfies the condition (2.4), if we select perturbation parameter ρ satisfying the condition (2.7), then for $X^{(k)}$ obtained by the expression (2.9), we have*

$$d^{(k)} < d^{(k-1)}, \quad k = 1, 2, \cdots, k_0 (\leq n)$$

and the number of elements in the set $M^{(k)}$ will be strictly greater than that of $M^{(k-1)}$, i.e. $M^{(k)} \supset M^{(k-1)}$, (or $p_k > p_{k-1}$).

Proof. see [3], Theorem 1.3.

Theorem 3. *There exists one and only one value ρ which satisfies the condition (2.7).*

Proof. see [3], Theorem 1.4.

Theorem 4: *Let $AX = Y$ be an overdetermined linear system of order $m \times n$ and suppose that Lemma 1 is satisfied up to $p_{k-1} > n$ or $k > n$ (note that this condition for A is weaker than the Haar condition which requires that arbitrary n row vectors of A are linearly independent), then $X^{(k)} = A^+ Y^{(k)}$ of (2.9) is the L_∞-solution of the system (2.1).*

Proof. see [3], Theorem 3.9.

Note. Suppose that Lemma 1 is satisfied up to $p_{k-1} = r$ or $k = r$. If $\operatorname{rank}(A) = r < n$, then $X^{(k)} = A^+ Y^{(k)}$ is the L_∞-solution which has minimal 2-norm.

For L_p-norm solution of the system (2.1) (steps $7° - 10°$ of the D.P.L.S-Algorithm) we have

Lemma 5. *$X^{(k)} = A^{(k)+} Y^{(k)}$ is the solution of the minimization problem*

$$\sum_{i=1}^{m} \frac{|r_i(X)|^2}{|r_i^{(k-1)}|^{2-p}} = \min_{X \in R^n} . \tag{3.4}$$

Proof. Because $A^{(k)+} = (\Sigma^{(k-1)} A)^+$ is the pseudo-inverse of the matrix A with weighting $(\Sigma^{(k-1)})^2$, so $X^{(k)} = A^{(k)+} Y^{(k)}$ is the minimal 2-norm solution of L.S-problem (3.4) with weighting $1/|r_i(X^{(k-1)})|^{2-p}$.

Theorem 6. *Let $X^{(k)}$ be an L_2-solution of the system (2.1), $r^{(k)} = Y - AX^{(k)} = r(X^{(k)})$ and assume that.*

(i) *if $\forall 1 \leq i \leq m$, $|r_i^{(k)}| > \varepsilon > 0$, let $X^{(k+1)}$ be a solution of the minimization* *problem*

$$\sum_{i=1}^{m} \frac{|r_i(X)|^{2Q}}{\sigma_i^{2Q-p}} = \min_{X\in R^n}, \quad \sigma_i = |r_i^{(k)}| \ , \ Q \in N \ , \ 0 < p \leq 2Q. \tag{3.5}$$

(ii) *if there exists a subset* $M_l = \{ j_1, \cdots, j_l \} \subseteq E$ *such that* $\forall\, j \in M_l$, $|r_j^{(k)}| \leq \varepsilon$, X_ε *is a solution of the problem*

$$\sum_{j=1}^{m} \frac{|r_j(X)|^{2Q}}{\sigma_j^{2Q-p}} = \min_{X\in R^n}, \quad \sigma_j = \begin{cases} \varepsilon, & j \in M_l, \\ |r_j^{(k)}|, & j \in E - M_l \end{cases} \tag{3.5'}$$

and $X^{(k+1)}$ *is a limit point of* $\{X_\varepsilon\}$ *as* $\varepsilon \longrightarrow 0$.

Then we have

1°

$$\Big(\sum_{i=1}^{m} |r_i^{(k+1)}|^p\Big)^{1/p} \leq \Big(\sum_{i=1}^{m} |r_i^{(k)}|^p\Big)^{1/p}. \tag{3.6}$$

If $AX = Y$ *has only a unique* L_2*-solution, then* $X^{(k)} = [A^{(k)}]^+ Y^{(k)}$ *is also unique for* $k = 0,\ 1,\ \cdots$, *and we have*

2°

$$\Big(\sum_{i=1}^{m} |r_i^{(k+1)}|^p\Big)^{1/p} < \Big(\sum_{i=1}^{m} |r_i^{(k)}|^p\Big)^{1/p}. \tag{3.7}$$

Proof. In the case (i), since $X^{(k+1)}$ is the solution of (3.5), so

$$\forall X \in R^n \ , \ \sum_{i=1}^{m} \frac{|r_i^{(k+1)}|^{2Q}}{|r_i^{(k)}|^{2Q-p}} \leq \sum_{i=1}^{m} \frac{|r_i(X)|^{2Q}}{|r_i^{(k)}|^{2Q-p}} \ .$$

The equality is achieved if and only if $X = X^{(k+1)}$. Set $X = X^{(k)}$. Then

$$\sum_{i=1}^{m} \frac{|r_i^{(k+1)}|^{2Q}}{|r_i^{(k)}|^{2Q-p}} < \sum_{i=1}^{m} |r_i^{(k)}|^p. \tag{3.8}$$

On the other hand, making use of the inequality of Hölder, we have

$$\sum_{i=1}^{m} |r_i^{(k+1)}|^p \leq \Big(\sum_{i=1}^{m} \frac{|r_i^{(k+1)}|^{2Q}}{|r_i^{(k)}|^{2Q-p}}\Big)^{p/2Q} \times \Big(\sum_{i=1}^{m} |r_i^{(k)}|^p\Big)^{(2Q-p)/2Q}.$$

From (3.8) we have

$$\sum_{i=1}^{m} |r_i^{(k+1)}|^p < \Big(\sum_{i=1}^{m} |r_i^{(k)}|^p\Big)^{p/2Q} \times \Big(\sum_{i=1}^{m} |r_i^{(k)}|^p\Big)^{(2Q-p)/2Q} = \sum_{i=1}^{m} |r_i^{(k)}|^p.$$

Then (3.7) is proved.

In the case (ii), without losing generality, we assume that $M_l = \{1\}$. Then (3.8) is replaced by

$$\frac{|r_1(X_\varepsilon)|^{2Q}}{\varepsilon^{2Q-p}} + \sum_{i=2}^{m} \frac{|r_i(X_\varepsilon)|^{2Q}}{|r_i^{(k)}|^{2Q-p}} < \varepsilon^p + \sum_{i=2}^{m} |r_i^{(k)}|^p. \tag{3.8'}$$

On the other hand, we have

$$\sum_{i=1}^{m} |r_i(X_\varepsilon)|^p \le \Big(\frac{|r_1(X_\varepsilon)|^{2Q}}{\varepsilon^{2Q-p}} + \sum_{i=2}^{m} \frac{|r_i(X_\varepsilon)|^{2Q}}{|r_i^{(k)}|^{2Q-p}}\Big)^{p/2Q} \times \Big(\varepsilon^p + \sum_{i=2}^{m} |r_i^{(k)}|^p\Big)^{(2Q-p)/2Q}$$
$$< \Big(\varepsilon^p + \sum_{i=2}^{m} |r_i^{(k)}|^p\Big)^{p/2Q} \times \Big(\varepsilon^p + \sum_{i=2}^{m} |r_i^{(k)}|^p\Big)^{(2Q-p)/2Q} = \varepsilon^p + \sum_{i=2}^{m} |r_i^{(k)}|^p \tag{3.8''}$$

It follows that there exists at least one limit point of $\{X_\varepsilon\}$ as $\varepsilon \longrightarrow 0$, denoted by $X^{(k+1)}$, and there exists a subsequence of $\{X_\varepsilon\}$ denoted by $\{X_{\varepsilon^{(l)}}\}_{l=1,2,\cdots}$ such that $\{X_{\varepsilon^{(l)}}\} \longrightarrow X^{(k+1)}$ as $\varepsilon^{(l)} \longrightarrow 0$.

From (3.8″) we have

$$\sum_{i=1}^{m} |r_i^{(k+1)}|^p = \lim_{\varepsilon^{(l)} \to 0} \sum_{i=1}^{m} |r(X_{\varepsilon^{(l)}})|^p < \lim_{\varepsilon^{(l)} \to 0} \Big(\sum_{i=2}^{m} |r_i^{(k)}|^p + (\varepsilon^{(l)})^p\Big)$$
$$= \sum_{i=2}^{m} |r_i^{(k)}|^p \le \sum_{i=1}^{m} |r_i^{(k)}|^p.$$

Then (3.7) is also proved.

Corollary 7. Let $Q = 1,\ 1 \le p < 2$, in Theorem 6. (3.7) may be written as

$$\|r(X^{(k+1)})\|_p < \|r(X^{(k)})\|_p. \tag{3.9}$$

Then $\{\|r(X^{(k)})\|_p\}$ is a strictly monotone decreasing bounded sequence (say $\|r^{(k)}\|_p > 0,\ \forall\ k$), with the limit $\tilde{m}$

$$\lim_{k \to \infty} \|r(X^{(k)})\|_p = \tilde{m}. \tag{3.10}$$

Because

$$\|r(X)\|_p = \Big(\sum_{i=1}^{m} |Y_i - \sum_{j=1}^{n} a_{ij} X_j|^p\Big)^{1/p}$$

is uniformly continuous and convex over R^n, we know that there exists at least a solution $\overline{X}$ of problem (1.4) such that ([2],p7)

$$\|r(\overline{X})\|_p = \min_{X \in R^n} \|r(X)\|_p.$$

From (2.10) and (3.10), we can know that

$$X^{(k)} = (\Sigma^{(k)}A)^{+}Y^{(k)} = (\Sigma^{(k)}A)^{+}\Sigma^{(k)}Y^{(0)}$$

is a sequence of bounded vectors, so we can make a closed sphere $S \subset R^n$ with the center $X^{(k)}$ as any approximate solution to $\overline{X}$ and a sufficiently large radius $\overline{d} > \|X^{(k)} - \overline{X}\|$ $(\forall k)$. Then

$$\forall k \quad X^{(k)} \in S, \quad \overline{X} \in S.$$

Now, because S is a compact convex set, the functional $\|r(X)\|_p$ has one and only one minimal value which can be achieved over S. When $k \longrightarrow \infty$, we have

$$\lim \|r(X^{(k)})\|_p = \|r(\overline{X})\|_p = \tilde{m}$$

and

$$\{X^{(k)}\} \longrightarrow \overline{X} \in S,$$

$$\lim_{k\to\infty} \sum_{i=1}^{m} \frac{|r_i(X^{(k+1)})|^2}{|r_i(X^{(k)})|^{2-p}} = \sum_{i=1}^{m} \frac{|r_i(\overline{X})|^2}{|r_i(\overline{X})|^{2-p}} = \sum_{i=1}^{m} |r_i(\overline{X})|^p = \min_{X\in R^n} \sum_{i=1}^{m} |r_i(X)|^p.$$

Consequently, $\overline{X}$ is one solution of problem (1.4).

§4. Numerical Examples

1. For data

X_i	0.0	1.0	2.0	3.0	4.0	5.0
Y_i	1.52	1.025	0.475	0.01	−0.475	−1.005

we require a fitting polynomial of order 1 ($Y = a + bX$) in the sense of L_2, L_∞, L_1 and $L_{1.5}$-norm. The results are given in Table 1.

2. Solve the system of overdetermined linear equations $Ax = Y, A \in R_2^{6\times 3}, Y \in R^6$, where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \\ 2 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} -3.0 \\ -1.0 \\ -7.0 \\ -11.1 \\ -6.9 \\ -7.2 \end{bmatrix}.$$

The solutions in $L_1, L_{1.5}, L_2, L_\infty$-norm and the corresponding error vectors are arranged in Table 2.

Table 1

Norm		L_2	L_∞	L_1	$L_{1.5}$
Solution	a	1.5147619048	1.5000000000	1.5200000160	1.5200054875
	b	−0.5025714286	−0.5000000000	−0.5033333515	−0.5038001340
	r_1	$0.52381 \times 10^{(-2)}$	$0.20000 \times 10^{(-1)}$	$-0.16040 \times 10^{(-7)}$	$-0.54875 \times 10^{(-5)}$
	r_2	$0.12810 \times 10^{(-1)}$	$+0.25000 \times 10^{(-1)}$	$0.83333 \times 10^{(-2)}$	$0.87946 \times 10^{(-2)}$
Residual	r_3	$-0.34619 \times 10^{(-1)}$	$-0.25000 \times 10^{(-1)}$	$-0.38333 \times 10^{(-1)}$	$-0.37405 \times 10^{(-1)}$
vector r	r_4	$0.29524 \times 10^{(-2)}$	$0.10000 \times 10^{(-1)}$	$0.38468 \times 10^{(-7)}$	$0.13949 \times 10^{(-2)}$
	r_5	$0.20524 \times 10^{(-1)}$	$+0.25000 \times 10^{(-1)}$	$0.18333 \times 10^{(-1)}$	$0.20195 \times 10^{(-1)}$
	r_6	$-0.69048 \times 10^{(-2)}$	$-0.05000 \times 10^{(-1)}$	$-0.83333 \times 10^{(-2)}$	$-0.60048 \times 10^{(-2)}$
$\|r(\ )\|_1$		0.83048×10^{-1}	0.11000	0.73333×10^{-1}	0.73800×10^{-1}
$\|r(\ )\|_2$		0.42316×10^{-1}	0.48990×10^{-1}	0.44096×10^{-1}	0.43844×10^{-1}
$\|r(\ )\|_\infty$		0.34619×10^{-1}	0.25000×10^{-1}	0.38333×10^{-1}	0.37405×10^{-1}
$\|r(\ )\|_{1.5}$					0.50790×10^{-1}

Table 2

Norm		L_2	L_∞	L_1	$L_{1.5}$
	x_1	−1.0370588235	−1.0000000000	−1.0451977992	−1.0441741413
Solution	x_2	−1.0370588235	−1.0000000000	−1.0451977992	−1.0441741413
	x_3	−1.8078431373	−2.0000000000	−1.7298022008	−1.7400826620
	r_1	0.88196	+1.00000	0.82080	0.82843
	r_2	−0.73373	−1.00000	−0.63941	−0.65173
Residual	r_3	−1.31020	−1.00000	−1.45000	−1.43149
vector	r_4	0.27961	0.90000	$0.17764 \times 10^{(-14)}$	$0.37027 \times 10^{(-1)}$
	r_5	−0.94392	−0.90000	−0.98941	−0.98322
	r_6	0.83020	0.80000	0.80099	0.80513
		L_2	L_∞	L_1	$L_{1.5}$
$\|r(\ )\|_1$		4.97961	5.60000	4.70000	4.73703
$\|r(\ )\|_2$		2.16592	2.29347	2.19193	2.18553
$\|r(\ )\|_\infty$		1.31020	1.00000	1.45000	1.43149
$\|r(\ )\|_{1.5}$					2.80480

References

[1] J.R. Rice, Numerical Methods, Software, and Analysis , McGraw-Hill, 1983.

[2] G.A. Watson, Approximate Theory and Numerical Methods, Wiley, 1980.

[3] Yang Shu-Guang, Algorithme de Perturbation Directrice pour le Probleme d'Ajustement des Donnees, *Acta Mathematica Scientia*, 8 (1988), 327-343.

[4] Yang Shu-Guang, Generalization of wilkinson iterative refinement-directional perturbation method, *Mathematica Applicata*, **2**:1 (1989), 37-46.