The Incremental Gauss-Newton Algorithm with Adaptive Stepsize Rule [1]

Hiroyuki Moriyama, Nobuo Yamashita and Masao Fukushima

Department of Applied Mathematics and Physics
Graduate School of Informatics
Kyoto University
Kyoto 606-8501, Japan

January 31, 2001

Abstract In this paper, we consider the Extended Kalman Filter (EKF) for solving nonlinear least squares problems. EKF is an incremental iterative method based on Gauss-Newton method that has nice convergence properties. Although EKF has the global convergence property under some conditions, the convergence rate is only sublinear under the same conditions. One of the reasons why EKF shows slow convergence is the lack of explicit stepsize. In the paper, we propose a stepsize rule for EKF and establish global convergence of the algorithm under the boundedness of the generated sequence and appropriate assumptions on the objective function. A notable feature of the stepsize rule is that the stepsize is kept greater than or equal to 1 at each iteration, and increases in the order $O(k)$. Therefore, we can expect that the proposed method converges faster than the original EKF. We report some numerical results, which demonstrate that the proposed method is promising.

1 Introduction

In this paper, we consider the nonlinear least squares problem of the form

$$\begin{array}{ll} \text{minimize} & f(x) = \dfrac{1}{2}\displaystyle\sum_{i=1}^{m} \|g_i(x)\|^2 \\ \text{subject to} & x \in \Re^n, \end{array} \tag{1}$$

where $g_i : \Re^n \to \Re^{r_i}, i = 1, \ldots, m$ are continuously differentiable functions. Least squares problems appear commonly in various fields such as curve fitting, neural network training, and pattern classification [2].

Let us denote $g(x) = [g_1(x)^T \cdots g_m(x)^T]^T$, and call $g_i(x)$ a data block and $g(x)$ the data set. Since $f(x)$ is equal to $\frac{1}{2}\|g(x)\|^2$, problem (1) is considered the problem of minimizing the residual of the system of equations $g(x) = 0$. Here $\|\cdot\|$ denotes the Euclidean norm of a vector, that is, $\|z\| = \sqrt{z^T z}$, where the superscript T denotes transposition. We define $f_i : \Re^n \to \Re$ by $f_i(x) = \frac{1}{2}\|g_i(x)\|^2$ for $i = 1, \ldots, m$. Then the objective function of problem (1) can be rewritten as $f(x) = \sum_{i=1}^{m} f_i(x)$. The gradient of f_i is given by $\nabla f_i(x) = \nabla g_i(x) g_i(x)$, where $\nabla g_i(x)$ denotes $n \times r_i$ Jacobian matrix of the function g_i.

A variety of methods have been proposed for solving least squares problems [2]. Roughly speaking, those methods are divided into two types, i.e., batch-type and incremental-type.

[1] This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan.

For the current iterate x^k, a batch-type algorithm generates the next iterate x^{k+1} by using information on all data blocks $g_i, i = 1, \ldots, m$. Batch-type algorithms contain the following well-known steepest descent method and Gauss-Newton method:

$$
\begin{array}{llcl}
\text{[Steepest descent method]} & x^{k+1} & := & x^k - \alpha^k \nabla f(x^k) \\
 & & = & x^k - \alpha^k \sum_{i=1}^m \nabla g_i(x^k) g_i(x^k), \\
\text{[Gauss-Newton method]} & x^{k+1} & := & x^k - \alpha^k (\nabla g(x^k) \nabla g(x^k)^T)^{-1} \nabla g(x^k) g(x^k),
\end{array}
$$

where α^k is a stepsize determined to decrease the objective function. Gauss-Newton method has such a desirable property as the quadratic rate of convergence under suitable assumptions. However, when the least squares problem has many data blocks, a batch-type algorithm may be ineffective because the size of the data set makes each iteration very costly. For such problems, it may be more effective to use an incremental-type algorithm explained below.

Incremental-type algorithms use information on a single data block g_i at each iteration. The following Back Propagation algorithm (BP for short) is a typical example of the incremental-type algorithms:

$$
\begin{array}{llcll}
\text{[BP]} & x_{i+1}^k & := & x_i^k - \alpha^k \nabla f_i(x_i^k), & i = 1, \ldots, m, \\
 & x_1^{k+1} & := & x_{m+1}^k, &
\end{array}
$$

BP was proposed as a method for solving the neural network training problem in [10] and has been studied extensively. It seems difficult that BP with a constant stepsize has a global convergence property. On the other hand, we can guarantee global convergence of BP with stepsize given by $\alpha^k = c/k$ at the kth iteration for some $c > 0$. But its convergence rate is only sublinear and very slow. Recently, Tseng [11] proposed BP using a new stepsize rule such that the stepsize does not tend to 0 but remains large. For the BP with the new stepsize rule, he proved global convergence and the linear rate of convergence under appropriate assumptions. However, in general, BP is not expected to converge quadratically because it is essentially based on the steepest descent method.

Of late years, some incremental algorithms based on Gauss-Newton method have been proposed from the viewpoint of accelerating convergence speed. The extended Kalman Filter (EKF for short) is one of such algorithms. EKF is an iterative method for solving least squares problems and has been used extensively in a variety of control applications. With an arbitrary initial point x_1^1, EKF generates a sequence $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ by

$$
\begin{array}{llcll}
\text{[EKF]} & x_{i+1}^k & := & x_i^k - d_i^k, & i = 1, \ldots, m, \\
 & x_1^{k+1} & := & x_{m+1}^k, &
\end{array} \tag{2}
$$

where d_i^k is a search direction given by

$$
d_i^k := (H_i^k)^{-1} \nabla g_i(x_i^k) g_i(x_i^k) \tag{3}
$$

and H_i^k is a symmetric positive-definite matrix successively determined by

$$
H_i^k := \lambda^k H_{i-1}^k + \nabla g_i(x_i^k) \nabla g_i(x_i^k)^T, \qquad i = 1, \ldots, m, \tag{4}
$$

$$
H_0^{k+1} := H_m^k, \tag{5}
$$

with $H_1^1 := \delta I$. Here δ is a positive scalar, λ^k is a parameter satisfying $0 < \lambda^k \le 1$, and I denotes the unit matrix.

In the case where $\lambda^k \equiv 1$ in (4), it is known that limit points of the sequence $\{x_1^k\}$ generated by EKF are stationary points of the least squares problem. However, the convergence rate is sublinear. On the other hand, when $\lambda^k \equiv \lambda < 1$, the generated sequence may converge to a certain point linearly, but the limit point is not necessarily a stationary point of problem (1).

To secure global convergence and a fast rate of convergence, Bertsekas [1] imposed the following condition on λ^k:

$$0 \leq 1 - (\lambda^k)^m \leq \frac{c}{k}, \tag{6}$$

where $c > 0$. Then, λ^k is small at an early stage of the iterations and approaches 1 as the iteration proceeds. He proved that EKF converges globally if λ^k satisfies (6). In the same paper, it was also shown that $\|(H_i^k)^{-1}\| = O(1/k)$. So, letting $(H_i^k)^{-1}$ play the role of a stepsize in (3), we can regard EKF as BP with stepsize of $O(1/k)$. This suggests that the convergence rate of EKF is sublinear, even if EKF is based on Gauss-Newton method. Such slow convergence is due to the lack of an explicit stepsize rule in EKF. Therefore, we may expect that the convergence rate can be improved by incorporating an efficient stepsize rule. In this paper, we propose an EKF algorithm with an explicit stepsize rule, so that the algorithm may retain global convergence while achieving a fast rate of convergence.

The EKF algorithm proposed in the paper is obtained by replacing the updating formula (2) with

$$x_{i+1}^k := x_i^k - \alpha^k d_i^k, \tag{7}$$

and (5) with

$$H_0^{k+1} := H_m^k + \delta^{k+1} I, \tag{8}$$

where δ^k is a nonnegative parameter. The reason why we use (8) is to guarantee the uniform positive definiteness of matrices H_i^k. Generally, a matrix $\nabla g_i(x) \nabla g_i(x)^T$ is positive semi-definite but not positive definite. Although the matrix H_i^k updated by (4) is nonsingular theoretically because of $H_1^1 = \delta I$, it may tend to be singular numerically. To keep matrices H_i^k uniformly positive definite for all $k = 1, 2, \ldots$, and any $i = 1, \ldots, m$, we replace the updating formula (5) with (8), and set $\delta^k > 0$, if necessary. We call the modified algorithm "EKF with stepsize". We give conditions under which the proposed algorithm is globally convergent, and evaluate its practical performance through numerical experiments.

The paper is organized as follows: In Section 2, we propose EKF with stepsize and show basic properties of the sequence generated by the algorithm. In Section 3, we give some assumptions and show some useful properties derived from them. By using those results, in Section 4, we establish global convergence of the proposed algorithm. In Section 5, we discuss how to implement the algorithm. Finally, we report numerical results in Section 6.

Throughout, superscript and subscript usually denote the iteration counts for the outer and the inner iterations, respectively. For a scalar λ, the ith power of λ is denoted by $(\lambda)^i$.

2 Algorithm

In this section, we propose the EKF with stepsize.

Extended Kalman Filter with stepsize (EKF-S)

Step.0 Choose the initial point $x_1^1 \in \Re^n$, parameters $c > 0$, $\alpha^1 > 1$, and $\delta^1 > 0$. Set $k := 1$, $i := 1$ and $H_0^1 := \delta^1 I$.

Step.1 Choose λ^k such that $0 \leq 1 - (\lambda^k)^m \leq c/k$.

Step.2 Obtain x_{i+1}^k from

$$x_{i+1}^k := x_i^k - \alpha^k (H_i^k)^{-1} \nabla g_i(x_i^k) g_i(x_i^k), \tag{9}$$

where

$$H_i^k := \lambda^k H_{i-1}^k + \nabla g_i(x_i^k) \nabla g_i(x_i^k)^T. \tag{10}$$

Step.3 If $i = m$, then go to **Step.4**. Otherwise, set $i := i + 1$ and go to **Step.2**.

Step.4 If a termination criterion is satisfied, then stop. Otherwise, choose $\delta^{k+1} > 0$ and set

$$x_1^{k+1} := x_{m+1}^k, \qquad H_0^{k+1} := H_m^k + \delta^{k+1} I. \tag{11}$$

Update α^{k+1}. Set $k := k + 1,\ i := 1$ and go to **Step.2**.

For the next iterate x_{i+1}^k obtained in Step.2, we have the following relation in a way similar to EKF [1]:

$$x_{i+1}^k = x_1^k - \alpha^k (H_i^k)^{-1} \sum_{j=1}^{i} (\lambda^k)^{i-j} \nabla g_j(x_j^k) \left(g_j(x_j^k) + \frac{1}{\alpha^k} \nabla g_j(x_j^k)^T (x_1^k - x_j^k) \right) \tag{12}$$

The proof of (12) is given later. The relation (12) plays an important role in the subsequent analysis. Here we discuss interpretation of (12). For the sake of brevity, let $\lambda^k \equiv 1$ and $\delta^k \equiv 0$, $\alpha^k = k$ and a generated sequence is bounded. Then, the term $\frac{1}{\alpha^k} \nabla g_j(x_j^k)^T (x_1^k - x_j^k)$ in (12) is negligible for sufficient large k. Thus, from (12), we can consider that the next iterate x_{i+1}^k is approximately obtained by

$$x_{i+1}^k \simeq x_1^k - \left(\frac{1}{k} H_i^k \right)^{-1} \sum_{j=1}^{i} \nabla g_j(x_j^k) g_j(x_j^k) \tag{13}$$

The matrix $\frac{1}{k} H_i^k$ in (13) can be viewed as an average of $\{\sum_{i=1}^m \nabla g_i(x_i^k) \nabla g_i(x_i^k)^T\}$. So, if the generated sequence converges to x^*, then $\{\frac{1}{k} H_i^k\}$ will converge to $\nabla g(x^*) \nabla g(x^*)^T$. Therefore, this algorithm is expected to behave as the Gauss-Newton method in a neighborhood of x^*. In fact, however, the global convergence of the algorithm may not be guaranteed in the case where $\alpha^k = k$. In the next section, we give the condition on α^k for its global convergence.

Now, we show the relation (12). To this end, we first consider the case where g_i are linear functions. The proof proceeds in a way similar to [1, 2].

Proposition 1 *Assume that B is an $n \times n$ symmetric positive definite matrix, and let $z_j \in \Re^{r_j}$ be given vectors and C_j be given $r_i \times n$ matrices, respectively, and let $0 < \lambda \leq 1$. Then, for $i = 1, \ldots, m$, the least squares estimates*

$$\psi_i = \arg \min_{x \in \Re^n} \left\{ (\lambda)^i (x - \psi_0)^T B (x - \psi_0) + \sum_{j=1}^{i} (\lambda)^{i-j} \| z_j - C_j x \|^2 \right\} \tag{14}$$

can be successively generated by the iterative scheme

$$\psi_i = \psi_{i-1} + (H_i)^{-1} C_i^T (z_i - C_i \psi_{i-1}), \quad i = 1, \ldots, m, \tag{15}$$

where ψ_0 is an arbitrary vector and H_i are positive-definite matrices generated by

$$H_i = \lambda H_{i-1} + C_i^T C_i, \quad i = 1, \ldots, m \tag{16}$$

with $H_0 = B$. More generally, for any $\bar{\imath}$, i such that $0 \le \bar{\imath} < i \le m$, we have

$$\psi_i = \psi_{\bar{\imath}} + (H_i)^{-1} \sum_{j=\bar{\imath}+1}^{i} (\lambda)^{i-j} C_j^T (z_j - C_j \psi_{\bar{\imath}}). \tag{17}$$

Before proving this proposition, we quote a result for the case where $m = 2$. By using the result, we prove Proposition 1.

Lemma 1 *[1, 2] Let ζ_1, ζ_2 be given vectors, and Γ_1, Γ_2 be given matrices such that $\Gamma_1^T\Gamma_1$ is positive definite. Then the vectors*

$$\tau_1 = \arg\min_{x\in\Re^n} \|\zeta_1 - \Gamma_1 x\|^2 \tag{18}$$

and

$$\tau_2 = \arg\min_{x\in\Re^n} \{\|\zeta_1 - \Gamma_1 x\|^2 + \|\zeta_2 - \Gamma_2 x\|^2\} \tag{19}$$

are given by

$$\tau_1 = \tau_0 + (\Gamma_1^T\Gamma_1)^{-1}\Gamma_1^T(\zeta_1 - \Gamma_1\tau_0) \tag{20}$$

and

$$\tau_2 = \tau_1 + (\Gamma_1^T\Gamma_1 + \Gamma_2^T\Gamma_2)^{-1}\Gamma_2^T(\zeta_2 - \Gamma_2\tau_1), \tag{21}$$

where τ_0 is an arbitrary vector.

Proof of Proposition 1: Consider two cases where $\bar{\imath} \ge 1$ and $\bar{\imath} = 0$.

(i) $0 < \bar{\imath} < i \le m$: Since B is symmetric positive definite, there exists a symmetric positive definite matrix $(B)^{\frac{1}{2}}$ such that $(B)^{\frac{1}{2}}(B)^{\frac{1}{2}} = B$. For $\psi_i, i = 1, \ldots, m$ given by (14), we have

$$\begin{aligned}
\psi_{\bar{\imath}} &= \arg\min_{x\in\Re^n} \Big\{(\lambda)^i(\psi_0 - x)^T B(\psi_0 - x) + \sum_{j=1}^{\bar{\imath}} (\lambda)^{i-j}\|z_j - C_j x\|^2\Big\} \\
&= \arg\min_{x\in\Re^n} \Big\{(\lambda)^i\|(B)^{\frac{1}{2}}(\psi_0 - x)\|^2 + \sum_{j=1}^{\bar{\imath}} (\lambda)^{i-j}\|z_j - C_j x\|^2\Big\}
\end{aligned}$$

and

$$\begin{aligned}
\psi_i &= \arg\min_{x\in\Re^n} \Big\{(\lambda)^i(\psi_0 - x)^T B(\psi_0 - x) + \sum_{j=1}^{i} (\lambda)^{i-j}\|z_j - C_j x\|^2\Big\} \\
&= \arg\min_{x\in\Re^n} \Big\{(\lambda)^i\|(B)^{\frac{1}{2}}(\psi_0 - x)\|^2 + \sum_{j=1}^{i} (\lambda)^{i-j}\|z_j - C_j x\|^2\Big\}.
\end{aligned} \tag{22}$$

Applying Lemma 1 with the correspondences $\tau_0 \sim \psi_0$, $\tau_1 \sim \psi_{\bar{\imath}}$, $\tau_2 \sim \psi_i$, and

$$\zeta_1 \sim \begin{pmatrix} \sqrt{(\lambda)^i}(B)^{\frac{1}{2}}\psi_0 \\ \sqrt{(\lambda)^{i-1}}z_1 \\ \vdots \\ \sqrt{(\lambda)^{i-\bar{\imath}}}z_{\bar{\imath}} \end{pmatrix}, \quad \Gamma_1 \sim \begin{pmatrix} \sqrt{(\lambda)^i}(B)^{\frac{1}{2}} \\ \sqrt{(\lambda)^{i-1}}C_1 \\ \vdots \\ \sqrt{(\lambda)^{i-\bar{\imath}}}C_{\bar{\imath}} \end{pmatrix},$$

$$\zeta_2 \sim \begin{pmatrix} \sqrt{(\lambda)^{i-\bar{\imath}-1}}z_{\bar{\imath}+1} \\ \vdots \\ z_i \end{pmatrix}, \quad \Gamma_2 \sim \begin{pmatrix} \sqrt{(\lambda)^{i-\bar{\imath}-1}}C_{\bar{\imath}+1} \\ \vdots \\ C_i \end{pmatrix},$$

we have from (20)

$$\begin{aligned}
\psi_{\bar\imath} &= \psi_0 + \Big((\lambda)^i B + \sum_{j=1}^{\bar\imath}(\lambda)^{i-j}C_j^T C_j\Big)^{-1}\sum_{j=1}^{\bar\imath}(\lambda)^{i-j}C_j^T(z_j - C_j\psi_0)\\
&= \psi_0 + (\lambda^{i-\bar\imath})^{-1}\Big((\lambda)^{\bar\imath} B + \sum_{j=1}^{\bar\imath}(\lambda)^{\bar\imath-j}C_j^T C_j\Big)^{-1}(\lambda^{i-\bar\imath})\sum_{j=1}^{\bar\imath}(\lambda)^{\bar\imath-j}C_j^T(z_j - C_j\psi_0)\\
&= \psi_0 + \Big((\lambda)^{\bar\imath} B + \sum_{j=1}^{\bar\imath}(\lambda)^{\bar\imath-j}C_j^T C_j\Big)^{-1}\sum_{j=1}^{\bar\imath}(\lambda)^{\bar\imath-j}C_j^T(z_j - C_j\psi_0),
\end{aligned}$$

and from (21)

$$\begin{aligned}
\psi_i &= \psi_{\bar\imath} + \Big(\Big((\lambda)^i B + \sum_{j=1}^{\bar\imath}(\lambda)^{i-j}C_j^T C_j\Big) + \Big(\sum_{j=\bar\imath+1}^{i}(\lambda)^{i-j}C_j^T C_j\Big)\Big)^{-1}\sum_{j=\bar\imath+1}^{i}(\lambda)^{i-j}C_j^T(z_j - C_j\psi_{\bar\imath})\\
&= \psi_{\bar\imath} + \Big((\lambda)^i B + \sum_{j=1}^{i}(\lambda)^{i-j}C_j^T C_j\Big)^{-1}\sum_{j=\bar\imath+1}^{i}(\lambda)^{i-j}C_j^T(z_j - C_j\psi_{\bar\imath}). \qquad (23)
\end{aligned}$$

Using (16) recursively, we obtain

$$\begin{aligned}
H_i &= \lambda H_{i-1} + C_i^T C_i\\
&= \lambda(\lambda H_{i-2} + C_{i-1}^T C_{i-1}) + C_i^T C_i\\
&\vdots\\
&= \lambda(\lambda\cdots(\lambda(\lambda B + C_1^T C_1) + C_2^T C_2) + \cdots + C_{i-1}^T C_{i-1}) + C_i^T C_i\\
&= (\lambda)^i B + \sum_{j=1}^{i}(\lambda)^{i-j}C_j^T C_j. \qquad (24)
\end{aligned}$$

By substituting (24) into (23), we have

$$\psi_i = \psi_{\bar\imath} + (H_i)^{-1}\sum_{j=\bar\imath+1}^{i}(\lambda)^{i-j}C_j^T(z_j - C_j\psi_{\bar\imath}), \quad 1 < \bar\imath < i \le m.$$

(ii) $\bar\imath = 0$: We can show this case by a similar argument, using the correspondences $\tau_0 \sim \psi_0$, $\tau_1 \sim \psi_{\bar\imath}$, $\tau_2 \sim \psi_i$, and

$$\zeta_1 \sim \sqrt{(\lambda)^i}(B)^{\frac12}\psi_0, \quad \Gamma_1 \sim \sqrt{(\lambda)^i}(B)^{\frac12}$$

$$\zeta_2 \sim \begin{pmatrix}\sqrt{(\lambda)^{i-1}}z_1\\ \sqrt{(\lambda)^{i-2}}z_2\\ \vdots\\ z_i\end{pmatrix}, \quad \Gamma_2 \sim \begin{pmatrix}\sqrt{(\lambda)^{i-1}}C_1\\ \sqrt{(\lambda)^{i-2}}C_2\\ \vdots\\ C_i\end{pmatrix}.$$

We obtain (15) as the special case of (17) with $\bar\imath = i-1$. **QED**

By using the proposition, we show the relation (12).

Lemma 2 *For a sequence $\{x_i^k\}$ generated by (9) and (10), we have*

$$x_{i+1}^k = x_1^k - \alpha^k(H_i^k)^{-1}\sum_{j=1}^{i}(\lambda^k)^{i-j}\nabla g_j(x_j^k)\left(g_j(x_j^k) + \frac{1}{\alpha^k}\nabla g_j(x_j^k)^T(x_1^k - x_j^k)\right). \qquad (25)$$

Proof : We will show (25) by using (17) in Proposition 1. First, note that the updating formula (9) of EKF-S can be rewritten as

$$\begin{aligned} x_{i+1}^k &:= x_i^k - \alpha^k (H_i^k)^{-1} \nabla g_i(x_i^k) g_i(x_i^k) \\ &= x_i^k - (H_i^k)^{-1} \nabla g_i(x_i^k) \left(\alpha^k g_i(x_i^k) \right). \end{aligned} \tag{26}$$

Consider the following correspondences

$$H_i \sim H_i^k, \quad z_j \sim \alpha^k g_j(\psi_{j-1}) - \nabla g_j(\psi_{j-1})^T \psi_{j-1}, \quad C_j \sim -\nabla g_j(\psi_{j-1})^T, \quad \psi_{j-1} \sim x_j^k.$$

Then, we see that (26) and (10) correspond to (15) and (16), respectively. Therefore, from Proposition 1, we obtain

$$\begin{aligned} x_{i+1}^k &= x_1^k - (H_i^k)^{-1} \sum_{j=1}^{i} (\lambda^k)^{i-j} \nabla g_j(x_j^k) \left(\alpha^k g_j(x_j^k) + \nabla g_j(x_j^k)^T (x_1^k - x_j^k) \right) \\ &= x_1^k - \alpha^k (H_i^k)^{-1} \sum_{j=1}^{i} (\lambda^k)^{i-j} \nabla g_j(x_j^k) \left(g_j(x_j^k) + \frac{1}{\alpha^k} \nabla g_j(x_j^k)^T (x_1^k - x_j^k) \right). \end{aligned}$$

QED

3 Assumptions and Lemmas

In this section, we give some assumptions needed for the global convergence of EKF-S, and show some lemmas derived from the assumptions.

The following are our assumptions.

Assumption 1 *A sequence* $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ *generated by EKF-S is contained in a compact set* X.

Assumption 2 *The functions* $f_i, i = 1, \ldots, m$ *are Lipschitz continuous on the set* X, *that is, there exists a scalar* $L > 0$ *such that*

$$\|\nabla f_i(x) - \nabla f_i(y)\| \le L \|x - y\|, \qquad \forall x, y \in X, \ \forall i = 1, \ldots, m. \tag{27}$$

Assumption 3 *The stepsize* α^k *used in EKF-S satisfies*

$$1 \le \alpha^k \le \max(1, \alpha_*^k), \quad k = 1, 2, \ldots, \tag{28}$$

where α_*^k *is given by*

$$\alpha_*^k = \begin{cases} \frac{(1-\eta)}{L} \frac{(x_1^{k+1} - x_1^k)^T H_m^k (x_1^{k+1} - x_1^k)}{\|x_1^{k+1} - x_1^k\| \sum_{i=2}^m \|x_i^k - x_1^k\| + \frac{m}{2} \|x_1^{k+1} - x_1^k\|^2}, & \text{if} \quad x_1^{k+1} \neq x_1^k \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

with $0 < \eta < 1$.

Assumption 4 *The parameters* δ^k *used in EKF-S are bounded above, that is, there exists a scalar* δ^U *such that* $0 \le \delta^k \le \delta^U$ *for all* k.

Assumption 5 *For the sequence* $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ *generated by EKF-S, the minimum eigenvalues of the matrices*

$$\sum_{j=1}^{i} \nabla g_j(x_j^{k+1}) \nabla g_j(x_j^{k+1})^T + \sum_{j=i+1}^{m} \nabla g_j(x_j^k) \nabla g_j(x_j^k)^T + \delta^{k+1} I$$

are greater than or equal to some constant $b > 0$ *for all* k *and* i.

We may expect that Assumption 2 usually holds for neural network training problems. Assumption 3 is a key condition on the stepsize α^k. Assumption 4 is needed to maintain the uniform positive definiteness of $\{H_i^k\}$.

To simplify the notation, we denote

$$D_i^k \equiv (H_i^k)^{-1}, \quad h_i^k \equiv \sum_{j=1}^{i} (\lambda^k)^{i-j} \nabla g_j(x_j^k) \Big(g_j(x_j^k) + \frac{1}{\alpha^k} \nabla g_j(x_j^k)^T (x_1^k - x_j^k) \Big). \tag{30}$$

Then (25) can be written as

$$x_{i+1}^k = x_1^k - \alpha^k D_i^k h_i^k. \tag{31}$$

Now we derive some lemmas from the assumptions, which will play an important role in showing the global convergence of EKF-S. By Assumptions 1, 2 and 3, we have the boundedness of $\{h_i^k\}$, $\{\nabla g_i(x_i^k) \nabla g_i(x_i^k)^T\}$, and $\{\nabla f_i(x_i^k)\}$.

Lemma 3 *(I) If Assumption 1 holds, there exists a scalar* $\rho_1 > 0$ *such that*

$$\|\nabla f_i(x)\| = \|\nabla g_i(x) g_i(x)\| \leq \rho_1, \qquad \forall x \in X, \ \forall i = 1, \ldots, m.$$

(II) If Assumption 1 holds, matrix $\nabla g_i(x) \nabla g_i(x)^T$ *is bounded on the set* X*, that is, there exists a scalar* $\rho_2 > 0$ *such that*

$$\|\nabla g_i(x) \nabla g_i(x)^T\| \leq \rho_2, \qquad \forall x \in X, \quad \forall i = 1, \ldots, m.$$

(III) If Assumptions 1, 2, and 3 hold, there exists a scalar $\rho_3 > 0$ *such that*

$$\|h_i^k\| < \rho_3, \qquad \forall k = 1, 2, \ldots, \ \forall i = 1, \ldots, m.$$

Proof :

(I) and (II). Since f_i and g_i are continuously differentiable by definition and X is bounded, the statements are true.

(III). From (30), we have for $i = 1, \ldots, m$

$$\begin{aligned}
\|h_i^k\| &= \left\| \sum_{j=1}^{i} (\lambda^k)^{i-j} \nabla g_j(x_j^k) \Big(g_j(x_j^k) + \frac{1}{\alpha^k} \nabla g_j(x_j^k)^T (x_1^k - x_j^k) \Big) \right\| \\
&\leq \sum_{j=1}^{i} \|\nabla g_j(x_j^k) g_j(x_j^k)\| + \frac{1}{\alpha^k} \sum_{j=1}^{i} \left\| \nabla g_j(x_i^k) \nabla g_j(x_j^k)^T (x_1^k - x_j^k) \right\| \\
&\leq i\rho_1 + \frac{\rho_2}{\alpha^k} \sum_{j=1}^{i} \|x_1^k - x_j^k\| \\
&\leq i\rho_1 + \rho_2 \sum_{j=1}^{i} \|x_1^k - x_j^k\|,
\end{aligned} \tag{32}$$

where the first inequality follows from $\lambda^k \leq 1$, the second inequality follows from (I) and (II), and the last inequality follows from Assumption 3. From Assumption 1, we see that $\sum_{j=1}^{i} \|x_1^k - x_i^k\|$ is bounded. Therefore, there exists a scalar $\rho_3 > 0$ such that $\|h_i^k\| < \rho_3$. **QED**

We need Assumption 5 to ensure the regularity of H_i^k.

Lemma 4 *Let Assumptions 1, 4, 5 hold. Let H_i^k be the matrices generated by (11) and D_i^k be their inverse matrices. Then, there exist positive scalars c_1 and c_2 such that for all k*

$$c_1 k\|x\|^2 \leq x^T H_i^k x \leq c_2 k\|x\|^2, \qquad \forall x \in \Re^n, \quad \forall i = 1, \ldots, m, \tag{33}$$

$$\frac{1}{c_2 k}\|x\|^2 \leq x^T D_i^k x \leq \frac{1}{c_1 k}\|x\|^2, \qquad \forall x \in \Re^n, \quad \forall i = 1, \ldots, m. \tag{34}$$

Proof : First, we show the relation between H_i^k and H_i^{k+1} for each $i = 1, \ldots, m$. Using (10) recursively, we have

$$\begin{aligned} H_i^{k+1} &= \lambda^{k+1} H_{i-1}^{k+1} + \nabla g_i(x_i^{k+1}) \nabla g_i(x_i^{k+1})^T \\ &= \lambda^{k+1}\left(\lambda^{k+1} H_{i-2}^{k+1} + \nabla g_{i-1}(x_{i-1}^{k+1}) \nabla g_{i-1}(x_{i-1}^{k+1})^T\right) + \nabla g_i(x_i^{k+1}) \nabla g_i(x_i^{k+1})^T \\ &\ \ \vdots \\ &= (\lambda^{k+1})^i H_0^{k+1} + \sum_{j=1}^{i} (\lambda^{k+1})^{i-j} \nabla g_j(x_j^{k+1}) \nabla g_j(x_j^{k+1})^T. \end{aligned} \tag{35}$$

Similarly, from (10) and (11) we have

$$\begin{aligned} H_0^{k+1} &= H_m^k + \delta^{k+1} I \\ &= \lambda^k H_{m-1}^k + \nabla g_m(x_m^k) \nabla g_m(x_m^k)^T + \delta^{k+1} I \\ &= \lambda^k\left(\lambda^k H_{m-2}^k + \nabla g_{m-1}(x_{m-1}^k) \nabla g_{m-1}(x_{m-1}^k)^T\right) + \nabla g_m(x_m^k) \nabla g_m(x_m^k)^T + \delta^{k+1} I \\ &\ \ \vdots \\ &= (\lambda^k)^{m-i} H_i^k + \sum_{j=i+1}^{m} (\lambda^k)^{m-j} \nabla g_j(x_j^k) \nabla g_j(x_j^k)^T + \delta^{k+1} I. \end{aligned} \tag{36}$$

Substituting (36) into (35) yields

$$\begin{aligned} H_i^{k+1} &= (\lambda^{k+1})^i \left\{ (\lambda^k)^{m-i} H_i^k + \sum_{j=i+1}^{m} (\lambda^k)^{m-j} \nabla g_j(x_j^k) \nabla g_j(x_j^k)^T + \delta^{k+1} I \right\} \\ &\quad + \sum_{j=1}^{i} (\lambda^{k+1})^{i-j} \nabla g_j(x_j^{k+1}) \nabla g_j(x_j^{k+1})^T \\ &= (\lambda^{k+1})^i (\lambda^k)^{m-i} H_i^k + \sum_{j=1}^{i} (\lambda^{k+1})^{i-j} \nabla g_j(x_j^{k+1}) \nabla g_j(x_j^{k+1})^T \\ &\quad + (\lambda^{k+1})^i \sum_{j=i+1}^{m} (\lambda^k)^{m-j} \nabla g_j(x_j^k) \nabla g_j(x_j^k)^T + (\lambda^{k+1})^i \delta^{k+1} I. \end{aligned} \tag{37}$$

Let ω_i^k be the largest eigenvalue of H_i^k. Since $\lambda^k \leq 1$ for all k, it follows from Lemma 3 (II) and (37) that

$$\begin{aligned}\omega_i^{k+1} &\leq \omega_i^k + i\rho_2 + (m-i)\rho_2 + \delta^{k+1} \\ &= \omega_i^k + m\rho_2 + \delta^{k+1}.\end{aligned}$$

By using this inequality recursively, we have

$$\begin{aligned}\omega_i^k &\leq (\omega_i^{k-2} + m\rho_2 + \delta^{k-1}) + m\rho_2 + \delta^k \\ &\vdots \\ &\leq \omega_i^1 + (k-1)m\rho_2 + \sum_{j=2}^{k} \delta^j. \end{aligned} \tag{38}$$

Letting $k = 0$ in (35), we have

$$\begin{aligned}H_i^1 &= (\lambda^1)^i H_0^1 + \sum_{j=1}^{i} (\lambda^1)^{i-j} \nabla g_j(x_j^1) \nabla g_j(x_j^1)^T \\ &= (\lambda^1)^i \delta^1 I + \sum_{j=1}^{i} (\lambda^1)^{i-j} \nabla g_j(x_j^1) \nabla g_j(x_j^1)^T.\end{aligned}$$

Since $\lambda^1 \leq 1$, the inequality $\omega_i^1 \leq \delta^1 + i\rho_2$ holds. Thus, we have from (38) we lead

$$\begin{aligned}\omega_i^k &\leq i\rho_2 + \delta^1 + (k-1)m\rho_2 + \sum_{j=2}^{k} \delta^j \\ &\leq km\rho_2 + \sum_{j=1}^{k} \delta^j, \end{aligned} \tag{39}$$

where the second inequality follows from $i \leq m$. Since there exists δ^U such that $\delta^j \leq \delta^U$ from Assumption 4, we obtain $\sum_{j=1}^{k} \delta^j \leq k\delta^U$. It follows from (39) that $\omega_i^k \leq kc_i^U$ for all k, where $c_i^U := m\rho_2 + \delta^U$.

Let v_i^k be the smallest eigenvalue of H_i^k, and let $\bar{k}$ be the smallest positive integer k such that $c/k < 1$. Now we consider the next two cases:

Case 1 : $k \geq \bar{k}$

In this case, we have from (6) that for each $i = 1, \ldots, m$,

$$(\lambda^{k+1})^i \geq (\lambda^{k+1})^m \geq 1 - \frac{c}{k+1} \geq 1 - \frac{c}{k} \tag{40}$$

$$(\lambda^k)^i \geq (\lambda^k)^m \geq 1 - \frac{c}{k}. \tag{41}$$

When $\lambda^{k+1} \geq \lambda^k$, we have from (40) and (41)

$$(\lambda^{k+1})^i (\lambda^k)^{m-i} \geq (\lambda^k)^m \geq 1 - \frac{c}{k},$$

We can also show this inequality similarly when $\lambda^k > \lambda^{k+1}$. So we have from (37)

$$v_i^{k+1} \geq \left(1 - \frac{c}{k}\right) v_i^k + \left(1 - \frac{c}{k}\right) b, \quad \forall k \geq 1. \tag{42}$$

Let

$$\bar{\beta}_i = \min\left\{\frac{(\bar{k}-c)b}{\bar{k}(1+c)}, \frac{v_i^{\bar{k}}}{\bar{k}}\right\}.$$

We will prove that $v_i^k \geq \bar{\beta}_i k,\ \forall k \geq \bar{k}$ by induction on k. Since $\bar{\beta}_i \leq v_i^{\bar{k}}/\bar{k}$, the inequality is true for some $\bar{k}$. Assume that $v_i^l \geq \bar{\beta}_i l$ holds for some $l \geq \bar{k}$. Then, we have

$$\begin{aligned}
\bar{\beta}_i \leq \frac{(\bar{k}-c)b}{(1+c)\bar{k}} &= \frac{b}{1+c} - \frac{cb}{(1+c)\bar{k}} \\
&\leq \frac{b}{1+c} - \frac{cb}{(1+c)l} \\
&= \frac{(l-c)b}{(1+c)l}.
\end{aligned}$$

Thus, we have

$$(c+1)\bar{\beta}_i \leq \left(1-\frac{c}{l}\right)b. \tag{43}$$

It then follows that

$$\begin{aligned}
v_i^{l+1} &\geq \left(1-\frac{c}{l}\right)\bar{\beta}_i l + \left(1-\frac{c}{l}\right)b \\
&\geq \bar{\beta}_i(l+1).
\end{aligned}$$

Consequently we obtain $v_i^k \geq \bar{\beta}_i k,\ \forall k \geq \bar{k}$.

Case 2 : $k < \bar{k}$.

Let

$$\hat{\beta}_i = \min_{k=1,\ldots,\bar{k}-1} \frac{v_i^k}{k}.$$

Then, it is clear that $v_i^k \geq \hat{\beta}_i k$ for all $k < \bar{k}$.

From the result of the two cases, setting c_i^L by

$$c_i^L = \min\left\{\bar{\beta}_i, \hat{\beta}_i\right\},$$

we see that $v_i^k \geq c_i^L k$ for all $k \geq 1$. Therefore, by taking $c_1 := \min_i c_i^L$ and $c_2 := \max_i c_i^U$, we see that the eigenvalue of the matrices H_i^k lie within the interval $[c_1 k, c_2 k]$, which shows (33). The inequality (34) directly follows from (33). **QED**

Next we show a property of α_*^k that appears in Assumption 3.

Lemma 5 *(a) Suppose that Assumptions 1, 4, and 5 hold, and let $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ be generated by EKF-S, If $x_i^k - x_1^k = 0,\ i = 2, \ldots, m+1$, then $h_i^k = 0,\ i = 1, \ldots, m$ and x_1^k is a stationary point of f.*

(b) Suppose that Assumptions 1, 4, and 5 hold. Then there exists a scalar ϕ such that $\alpha_*^k \leq \phi k$ for all k.

Proof :

(a) From (31) we have $x_i^k - x_1^k = \alpha^k D_i^k h_i^k = 0$. It follows from the positive definiteness of D_i^k and (30) that $h_i^k = 0, i = 1, \ldots, m$. Furthermore it is clear that $x_1^k = x_2^k = \cdots = x_{m+1}^k \equiv x_1^{k+1}$. Thus, from (9), we have $\alpha^k (H_i^k)^{-1} \nabla f_i(x_i^k) = 0, i = 1, \ldots, m$. Taking account of Assumption 3 and Lemma 4, it is seen that $\nabla f_i(x_i^k) = 0, i = 1, \ldots, m$. Therefore, we have $\nabla f_i(x_1^k) = 0, i = 1, \ldots, m$, which means $\nabla f(x_1^k) = 0$.

(b) If $\|h_m^k\| \neq 0$, then from (29), (31) and Lemma 4, we have

$$
\begin{aligned}
\alpha_*^k &= \frac{1-\eta}{L} \frac{(h_m^k)^T D_m^k h_m^k}{\|D_m^k h_m^k\| \sum_{i=2}^m \|D_{i-1}^k h_{i-1}^k\| + \frac{m}{2}\|D_m^k h_m^k\|^2} \\
&\leq \frac{1-\eta}{L} \frac{(c_1 k)^2}{c_2 k} \frac{\|h_m^k\|^2}{\|h_m^k\| \sum_{i=2}^m \|h_{i-1}^k\| + \frac{m}{2}\|h_m^k\|^2} \\
&= \frac{1-\eta}{L} \frac{(c_1)^2}{c_2} k \frac{1}{(\sum_{i=2}^m \|h_{i-1}^k\|)/\|h_m^k\| + \frac{m}{2}} \\
&\leq \frac{2(c_1)^2(1-\eta)}{c_2 mL} k
\end{aligned}
$$

By taking $\phi = \frac{2(c_2)^2(1-\eta)}{(c_1)^2 mL}$, we obtain $\alpha_*^k \leq \phi k$ for all k. If $\|h_m^k\| = 0$, from Assumption 3, we have $\alpha_*^k = 0$. Then, it is clear that $\alpha_*^k \leq \phi k$. The proof is completed. **QED**

4 Global Convergence Analysis

In this section, under Assumptions 1–5, we show that EKF has the global convergence property. We note that if $\sum_{i=2}^{m+1} \|x_i^k - x_1^k\| = 0$, then x_i^k is a stationary point of the least squares problem by Lemma 5. Therefore, we suppose that $\sum_{i=2}^{m+1} \|x_i^k - x_1^k\| \neq 0$ for all k in the subsequent analysis.

First, we consider the behavior of $\{h_m^k\}$ as $k \to \infty$.

Lemma 6 *Let Assumptions 1, 2, 3, 4 and 5, and let the sequence* $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ *be generated by EKF-S. Then, we have* $\liminf_{k\to\infty} \|h_m^k\| = 0$.

Proof : Since $\nabla f = \nabla f_1 + \cdots + \nabla f_m$ is Lipschitz continuous with constant mL on X by Assumptions 1 and 2, it follows from [2, Descent Lemma] that for any k

$$
\begin{aligned}
f(x_1^{k+1}) - f(x_1^k) &\leq \langle \nabla f(x_1^k), x_1^{k+1} - x_1^k \rangle + \frac{mL}{2}\|x_1^{k+1} - x_1^k\|^2 \\
&= \langle \nabla f(x_1^k) - h_m^k, x_1^{k+1} - x_1^k \rangle + \langle h_m^k, x_1^{k+1} - x_1^k \rangle + \frac{mL}{2}\|x_1^{k+1} - x_1^k\|^2 \\
&\leq \|\nabla f(x_1^k) - h_m^k\| \|x_1^{k+1} - x_1^k\| + \langle h_m^k, x_1^{k+1} - x_1^k \rangle + \frac{mL}{2}\|x_1^{k+1} - x_1^k\|^2 \qquad (44)
\end{aligned}
$$

Moreover, from (30) we have

$$
\begin{aligned}
&\nabla f(x_1^k) - h_m^k \\
&= \sum_{i=1}^m \nabla g_i(x_1^k) g_i(x_1^k) - \sum_{i=1}^m (\lambda^k)^{m-i} \nabla g_i(x_i^k) \Big(g_i(x_i^k) + \frac{1}{\alpha^k} \nabla g_i(x_i^k)^T (x_1^k - x_i^k) \Big) \\
&= \sum_{i=1}^m \Big(1 - (\lambda^k)^{m-i}\Big) \nabla g_i(x_1^k) g_i(x_1^k) + \sum_{i=1}^m (\lambda^k)^{m-i} \Big\{ \nabla g_i(x_1^k) g_i(x_1^k) - \nabla g_i(x_i^k) g_i(x_i^k) \Big\}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1}{\alpha^k}\sum_{i=1}^{m}(\lambda^k)^{m-i}\left\{\nabla g_i(x_i^k)\nabla g_i(x_i^k)^T(x_1^k - x_i^k)\right\}\\
= \;& \sum_{i=1}^{m}\left(1-(\lambda^k)^{m-i}\right)\nabla f_i(x_1^k) + \sum_{i=2}^{m}(\lambda^k)^{m-i}\left\{\nabla f_i(x_1^k) - \nabla f_i(x_i^k)\right\}\\
&-\frac{1}{\alpha^k}\sum_{i=2}^{m}(\lambda^k)^{m-i}\left\{\nabla g_i(x_i^k)\nabla g_i(x_i^k)^T(x_1^k - x_i^k)\right\}.
\end{aligned}
$$

It then follows from $\lambda^k \le 1$ that

$$
\begin{aligned}
&\|\nabla f(x_1^k) - h_m^k\|\\
\le\;& \left(1-(\lambda^k)^m\right)\sum_{i=1}^{m}\|\nabla f_i(x_1^k)\| + \sum_{i=2}^{m}\|\nabla f_i(x_1^k) - \nabla f_i(x_i^k)\|\\
&+\frac{1}{\alpha^k}\sum_{i=2}^{m}\|\nabla g_i(x_i^k)\nabla g_i(x_i^k)^T\|\|x_1^k - x_i^k\|\\
\le\;& \left(1-(\lambda^k)^m\right)\sum_{i=1}^{m}\|\nabla f_i(x_1^k)\| + L\sum_{i=2}^{m}\|x_1^k - x_i^k\| + \frac{\rho_2}{\alpha^k}\sum_{i=2}^{m}\|x_1^k - x_i^k\|\\
=\;& \left(1-(\lambda^k)^m\right)\sum_{i=1}^{m}\|\nabla f_i(x_1^k)\| + (L + \frac{\rho_2}{\alpha^k})\sum_{i=2}^{m}\|x_1^k - x_i^k\|, \qquad (45)
\end{aligned}
$$

where the second inequality follows from Assumption 2. Note that $x_i^k - x_1^k = -\alpha^k D_{i-1}^k h_{i-1}^k$ by (31), and hence $x_1^{k+1} - x_1^k = x_{m+1}^k - x_1^k = -\alpha^k D_m^k h_m^k$. It then follows from (44) and (45) that

$$
\begin{aligned}
&f(x_1^{k+1}) - f(x_1^k)\\
\le\;& \left\{\left(1-(\lambda^k)^m\right)\sum_{i=1}^{m}\|\nabla f_i(x_1^k)\| + (L+\frac{\rho_2}{\alpha^k})\sum_{i=2}^{m}\|x_1^k - x_i^k\|\right\}\|x_1^{k+1} - x_1^k\|\\
&-\alpha^k(h_m^k)^T D_m^k h_m^k + \frac{mL}{2}\|x_1^{k+1} - x_1^k\|^2\\
\le\;& \left\{(1-(\lambda^k)^m)\sum_{i=1}^{m}\|\nabla f_i(x_1^k)\| + (\alpha^k L + \rho_2)\sum_{i=2}^{m}\|D_{i-1}^k h_{i-1}^k\|\right\}\alpha^k\|D_m^k h_m^k\|\\
&-\alpha^k(h_m^k)^T D_m^k h_m^k + \frac{mL}{2}\left(\alpha^k\|D_m^k h_m^k\|\right)^2. \qquad (46)
\end{aligned}
$$

Consider the two case where (i) $\alpha^k = 1$ and (ii) $1 < \alpha^k \le \alpha_*^k$ in (46). For convenience, we define two index sets K_1 and K_2 by

$$
K_1 = \{k \mid \alpha^k = 1\}, \qquad K_2 = \{k \mid 1 < \alpha^k \le \alpha_*^k\}.
$$

Since $1 \le \alpha^k \le \max(1, \alpha_*^k)$, it is easy to see that $K_1 \cap K_2 = \emptyset$ and $K_1 \cup K_2 = \{1, 2, \ldots\}$.

First, we consider case (i). We have

$$
\begin{aligned}
&f(x_1^{k+1}) - f(x_1^k)\\
\le\;& (1-(\lambda^k)^m)\|D_m^k h_m^k\|\sum_{i=1}^{m}\|\nabla f_i(x_1^k)\| + (L+\rho_2)\|D_m^k h_m^k\|\sum_{i=2}^{m}\|D_{i-1}^k h_{i-1}^k\|\\
&-(h_m^k)^T D_m^k h_m^k + \frac{mL}{2}\|D_m^k h_m^k\|^2
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{c}{k}\frac{1}{c_1 k}\|h_m^k\|\sum_{i=1}^m \|\nabla f_i(x_1^k)\| + \left(\frac{1}{c_1 k}\right)^2 (L+\rho_2)\|h_m^k\|\sum_{i=2}^m \|h_{i-1}^k\| \\
&\quad -\frac{1}{c_2 k}\|h_m^k\|^2 + \frac{mL}{2}\left(\frac{1}{c_1 k}\right)^2 \|h_m^k\|^2 \\
&= -\frac{1}{c_2 k}\|h_m^k\|^2 + \frac{\|h_m^k\|}{(k)^2}\Big\{\frac{c}{c_1}\sum_{i=1}^m \|\nabla f_i(x_1^k)\| + \frac{L+\rho_2}{(c_1)^2}\sum_{i=2}^m \|h_{i-1}^k\| + \frac{mL}{2(c_1)^2}\|h_m^k\|\Big\}, \qquad (47)
\end{aligned}
$$

where the second inequality follows from (34). Moreover, since $\{\|h_i^k\|\}$ and $\{\|\nabla f_i(x_1^k)\|\}$ are bounded by Lemma 3, there exist positive scalars ζ_1 and ζ_2 such that

$$
f(x_1^{k+1}) \le f(x_1^k) - \frac{\zeta_1}{k}\|h_m^k\|^2 + \frac{\zeta_2}{(k)^2}\|h_m^k\|. \qquad (48)
$$

Next, we consider case (ii) $1 < \alpha^k \le \alpha_*^k$, that is, $k \in K_2$. From the definition (29) of α_*^k, we have

$$
\begin{aligned}
\alpha^k \le \alpha_*^k &= \frac{(1-\eta)}{L}\frac{(x_1^{k+1}-x_1^k)^T H_m^k (x_1^{k+1}-x_1^k)}{\|x_1^{k+1}-x_1^k\|\sum_{i=2}^m \|x_{i-1}^k - x_1^k\| + \frac{m}{2}\|x_1^{k+1}-x_1^k\|^2} \\
&= \frac{1-\eta}{L}\frac{(h_m^k)^T D_m^k h_m^k}{\|D_m^k h_m^k\|\sum_{i=2}^m \|D_{i-1}^k h_{i-1}^k\| + \frac{m}{2}\|D_m^k h_m^k\|^2}. \qquad (49)
\end{aligned}
$$

Rearranging this inequality yields

$$
(\alpha^k)^2 L\|D_m^k h_m^k\|\sum_{i=2}^m \|D_{i-1}^k h_{i-1}^k\| + \frac{mL}{2}(\alpha^k\|D_m^k h_m^k\|)^2 - \alpha^k (h_m^k)^T D_m^k h_m^k \le -\eta\alpha^k (h_m^k)^T D_m^k h_m^k.
$$

Substituting it into (46), we have

$$
\begin{aligned}
&f(x_1^{k+1}) - f(x_1^k) \\
&\le -\eta\, \alpha^k (h_m^k)^T D_m^k h_m^k + \alpha^k\left(1-(\lambda^k)^m\right)\|D_m^k h_m^k\|\sum_{i=1}^m \|\nabla f_i(x_1^k)\| + \alpha^k \rho_2 \|D_m^k h_m^k\|\sum_{i=2}^m \|D_{i-1}^k h_{i-1}^k\| \\
&\le -\eta\frac{1}{c_2 k}\alpha^k\|h_m^k\|^2 + \frac{c}{k}\frac{1}{c_1 k}\alpha^k\|h_m^k\|\sum_{i=1}^m \|\nabla f_i(x_1^k)\| + \left(\frac{1}{c_1 k}\right)^2 \rho_2 \alpha^k \|h_m^k\|\sum_{i=2}^m \|h_{i-1}^k\| \\
&= -\frac{\eta}{c_2}\frac{1}{k}\alpha^k\|h_m^k\|^2 + \frac{1}{(k)^2}\alpha^k\|h_m^k\|\left\{\frac{c}{c_1}\sum_{i=1}^m \|\nabla f_i(x_1^k)\| + \frac{\rho_2}{(c_1)^2}\sum_{i=2}^m \|h_{i-1}^k\|\right\}, \qquad (50)
\end{aligned}
$$

where the second inequality follows from (34). Therefore, since $\{\|\nabla f_i(x_1^k)\|\}$ and $\{\|h_i^k\|\}$ are bounded by Lemma 3, there exist scalars $\xi > 0$ and $\xi_2 > 0$ such that

$$
f(x_1^{k+1}) \le f(x_1^k) - \frac{\xi_1}{k}\alpha^k\|h_m^k\|^2 + \frac{\xi_2}{(k)^2}\alpha^k\|h_m^k\|. \qquad (51)
$$

Now, let $\bar{k}$ be an arbitrary integer. For $k > \bar{k}$, we define $\Lambda = \{\bar{k}, \dots, k\}$. Summing (48) and (51) over all $l \in \Lambda$, we obtain

$$
\begin{aligned}
f(x_1^{k+1}) - f(x_1^{\bar{k}}) \;\le\; & -\xi_1 \sum_{l \in K_1 \cap \Lambda} \frac{\|h_m^l\|^2}{l} + \xi_2 \sum_{l \in K_1 \cap \Lambda} \frac{\|h_m^k\|}{(l)^2} \\
& -\zeta_1 \sum_{l \in K_2 \cap \Lambda} \frac{\alpha^l\|h_m^l\|^2}{l} + \zeta_2 \sum_{l \in K_2 \cap \Lambda} \frac{\alpha^l\|h_m^k\|}{(l)^2}.
\end{aligned}
$$

Taking $\sigma_1 := \min(\zeta_1, \xi_1) > 0$ and $\sigma_2 := \max(\zeta_2, \xi_2) > 0$ and noting that $\alpha^l = 1$ for all $l \in K_1 \cap \Lambda$, we have

$$\begin{aligned} f(x_1^{k+1}) &\leq f(x_1^{\bar{k}}) - \sigma_1 \sum_{l=\bar{k}}^{k} \frac{\alpha^l \|h_m^l\|^2}{l} + \sigma_2 \sum_{l=\bar{k}}^{k} \frac{\alpha^l \|h_m^k\|}{(l)^2} &(52)\\ &= f(x_1^{\bar{k}}) - \sum_{l=\bar{k}}^{k} \frac{1}{l} \sigma_1 \alpha^l \|h_m^l\| \Big(\|h_m^l\| - \frac{\sigma_2}{\sigma_1} \frac{1}{l} \Big). &(53) \end{aligned}$$

Since $\{f(x_1^k)\}$ is bounded below and $\sum_{l=\bar{k}}^{\infty} \frac{1}{l} = \infty$, there cannot exist a constant $e > 0$ such that $\alpha^k \|h_m^k\| (\|h_m^k\| - \frac{\sigma_2}{\sigma_1} \frac{1}{k}) > e, \forall k \geq \bar{k}$. So we obtain

$$\liminf_{k \to \infty} \alpha^k \|h_m^k\| (\|h_m^k\| - \frac{\sigma_2}{\sigma_1} \frac{1}{k}) \leq 0.$$

Since $\lim_{k \to \infty} \frac{1}{k} = 0$ and $\alpha^k \geq 1$ for all k, it follows that there cannot exist an $e > 0$ such that $\|h_m^k\| > e, \forall k \geq \bar{k}$. Therefore we have $\liminf_{k \to \infty} \|h_m^k\| = 0$. **QED**

Next, we show the relation between h_m^k and $\nabla f(x_1^k)$.

Lemma 7 *Let Assumptions 1, 2, 3, 4, and 5 hold and let the sequence* $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ *be generated by EKF-S. Then we have*

$$\lim_{\substack{k \to \infty \\ k \in \mathcal{D}}} \|\nabla f(x_1^k)\| = 0,$$

where $\mathcal{D}$ *is a set of integers such that* $\{\|h_m^k\|\}_{k \in \mathcal{D}} \to 0$*. Therefore, if* $\liminf_{k \to \infty} \|h_m^k\| = 0$*, then we have* $\liminf_{k \to \infty} \|\nabla f(x_1^k)\| = 0$*.*

Proof : From (45), we have

$$\|\nabla f(x_1^k) - h_m^k\| \leq \Big(1 - (\lambda^k)^m\Big) \sum_{i=1}^{m} \|\nabla f_i(x_1^k)\| + \Big(L + \frac{\rho_2}{\alpha^k}\Big) \sum_{i=2}^{m} \|x_i^k - x_1^k\|.$$

Since $\lambda^k \to 1$, $\alpha^k \geq 1$ and $\|\nabla f_i(x_1^k)\|$ is bounded, it suffices to show $\lim_{k \to \infty, k \in \mathcal{D}} \sum_{i=2}^{m} \|x_i^k - x_1^k\| = 0$. In a way similar to the proof of Lemma 6, we define K_1 and K_2 by

$$K_1 = \{k \mid \alpha^k = 1\}, \qquad K_2 = \{k \mid 1 < \alpha^k \leq \alpha_*^k\}.$$

When $\alpha^k = 1$, from (31) we have

$$\begin{aligned} \sum_{i=2}^{m} \|x_i^k - x_1^k\| &= \sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| \\ &\leq \frac{1}{c_1 k} \sum_{i=2}^{m} \|h_{i-1}^k\|, &(54) \end{aligned}$$

where the inequality follows from (34). Since $\{\|h_i^k\|\}$ is bounded from Lemma 3, it follows from (54) that

$$\lim_{k \to \infty, k \in \mathcal{D} \cap K_1} \sum_{i=2}^{m} \|x_i^k - x_1^k\| = 0.$$

When $1 < \alpha^k \le \alpha_*^k$, from (49) we have

$$\begin{aligned}
\sum_{i=1}^{m} \|x_i^k - x_1^k\| &= \sum_{i=2}^{m} \alpha^k \|D_{i-1}^k h_{i-1}^k\| \\
&\le \frac{1-\eta}{L} \frac{(h_m^k)^T D_m^k h_m^k}{\|D_m^k h_m^k\| \sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| + \frac{m}{2}\|D_m^k h_m^k\|^2} \sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| \\
&\le \frac{1-\eta}{L} \frac{\|h_m^k\| \|D_m^k h_m^k\|}{\|D_m^k h_m^k\| \sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| + \frac{m}{2}\|D_m^k h_m^k\|^2} \sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| \\
&= \frac{1-\eta}{L} \frac{\|h_m^k\|}{\sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| + \frac{m}{2}\|D_m^k h_m^k\|} \sum_{i=2}^{m} \|D_{i-1}^k h_{i-1}^k\| \\
&\le \frac{1-\eta}{L} \frac{c_2 k}{c_1 k} \frac{\|h_m^k\| \sum_{i=2}^{m} \|h_{i-1}^k\|}{\sum_{i=2}^{m} \|h_{i-1}^k\| + \frac{m}{2}\|h_m^k\|} \\
&= \frac{1-\eta}{L} \frac{c_2}{c_1} \frac{\|h_m^k\| \sum_{i=2}^{m} \|h_{i-1}^k\|}{\sum_{i=2}^{m} \|h_{i-1}^k\| + \frac{m}{2}\|h_m^k\|}.
\end{aligned} \tag{55}$$

First we consider the case where $\liminf_{k\to\infty, k\in\mathcal{D}\cap K_2} \sum_{i=2}^{m} \|h_{i-1}^k\| = 0$. Since $\lim_{k\to\infty, k\in\mathcal{D}\cap K_2} \|h_m^k\| = 0$, it is easy to see that $\lim_{k\to\infty, k\in\mathcal{D}\cap K_2} \sum_{i=2}^{m} \|x_1^k - x_i^k\| = 0$. Next, we consider the case where $\liminf_{k\to\infty, k\in\mathcal{D}\cap K_2} \sum_{i=2}^{m} \|h_{i-1}^k\| > 0$. It is also clear that the same conclusion is obtained. Thus, we have

$$\lim_{k\to\infty, k\in\mathcal{D}} \sum_{i=2}^{m} \|x_i^k - x_1^k\| = 0$$

Therefore, the proof is completed. **QED**

Note that Lemmas 6 and 7 yield $\liminf_{k\to\infty} \|\nabla f(x_1^k)\| = 0$.

Next, we show the that $\{f(x_1^k)\}_{k=1,2,\dots}$ converges.

Lemma 8 *Let Assumptions 1, 2, 3, 4, and 5 hold, and let the sequence $\{x_1^k, \dots, x_m^k\}_{k=1,2,\dots}$ be generated by EKF-S. Then the sequence $\{f(x_1^k)\}_{k=1,2,\dots}$ converges.*

Proof : From (53), we have that for an arbitrary $k \ge \bar{k}$

$$f(x_1^{k+1}) \le f(x_1^{\bar{k}}) - \sum_{l=\bar{k}}^{k} \frac{1}{l} \sigma_1 \alpha^l \|h_m^l\| \Big(\|h_m^l\| - \frac{\sigma_2}{\sigma_1} \frac{1}{l} \Big).$$

Here we define index sets $\mathcal{K}_1$ and $\mathcal{K}_2$ by

$$\mathcal{K}_1 = \Big\{ k \;\Big|\; \|h_m^k\| - \frac{\sigma_2}{\sigma_1}\frac{1}{k} > 0,\ k \ge \bar{k} \Big\},$$

$$\mathcal{K}_2 = \Big\{ k \;\Big|\; \|h_m^k\| - \frac{\sigma_2}{\sigma_1}\frac{1}{k} \le 0,\ k \ge \bar{k} \Big\}.$$

For $k \in \mathcal{K}_1$, we have

$$\begin{aligned}
f(x_1^{k+1}) - f(x_1^k) &\le -\sigma_1 \frac{\alpha^k}{k} \|h_m^k\| (\|h_m^k\| - \frac{\sigma_2}{\sigma_1}\frac{1}{k}) \\
&\le 0.
\end{aligned} \tag{56}$$

For $k \in \mathcal{K}_2$, we have

$$
\begin{aligned}
f(x_1^{k+1}) - f(x_1^k) &\leq \sigma_1 \frac{\alpha^k}{k} \|h_m^k\| \Big(\frac{\sigma_2}{\sigma_1} \frac{1}{k} - \|h_m^k\| \Big) \\
&\leq \sigma_1 \frac{\alpha^k}{k} \|h_m^k\| \Big(\frac{\sigma_2}{\sigma_1} \frac{1}{k} \Big) \\
&\leq \sigma_1 \frac{\alpha^k}{k} \Big(\frac{\sigma_2}{\sigma_1} \frac{1}{k} \Big) \Big(\frac{\sigma_2}{\sigma_1} \frac{1}{k} \Big) \\
&\leq \frac{(\sigma_2)^2}{\sigma_1} \frac{\phi}{(k)^2},
\end{aligned}
$$

where the third inequality follows from the fact that $\|h_m^k\| \leq \frac{\sigma_2}{\sigma_1} \frac{1}{k}$, the last inequality follows from Lemma 5 and $1 \leq \alpha^k \leq \max(1, \alpha_*^k)$, and ϕ is the constant indicated in Lemma 5. Let

$$
\nu^k = f(x_1^k), \qquad \mu^k = \begin{cases} 0, & k \in \mathcal{K}_1 \\ \frac{(\sigma_2)^2 \phi}{\sigma_1} \frac{1}{(k)^2}, & k \in \mathcal{K}_2 \end{cases}.
$$

Then, $\{\nu^k\}$ and $\{\mu^k\}$ are nonnegative sequences such that $\nu^{k+1} \leq \nu^k + \mu^k, \forall k$ and $\sum_{k=1}^{\infty} \mu^k < \infty$. It then follows from [9, supermartingale convergence theorem] that $\{\nu^k\}$ converges, which means that $\{f(x_1^k)\}_{k=1,2,\ldots}$ converges. **QED**

Using the preceding results, we can prove the following global convergence theorem.

Theorem 1 *Suppose that Assumption 1, 2, 3, 4, and 5 hold, and let the sequence $\{x_1^k, \ldots, x_m^k\}_{k=1,2,\ldots}$ be generated by EKF-S. Then we have*

$$
\lim_{k \to \infty} \|\nabla f(x_1^k)\| = 0.
$$

Consequently, all the cluster points of the sequence $\{x_1^k\}_{k=1,2,\ldots}$ are stationary points of the problem (1).

Proof : We assume that $\limsup_{k \to \infty} \|\nabla f(x_1^k)\| > 0$ for the purpose of contraction. Since $\limsup_{k \to \infty} \|\nabla f(x_1^k)\| > 0$, there exists a positive scalar ε such that the set

$$
\mathcal{K} = \{ \ k \mid \|\nabla f(x_1^k)\| > \varepsilon \} \neq \emptyset \tag{57}
$$

has an infinite number of elements. For each $k \in \mathcal{K}$, define $l(k)$ by

$$
l(k) = \min\{ l > k \mid \|\nabla f(x_1^l)\| < \frac{\varepsilon}{2} \}. \tag{58}
$$

Note that for all $k \in \mathcal{K}$,

$$
\|\nabla f(x_1^l)\| > \frac{\varepsilon}{2}, \qquad \forall l = k, k+1, \ldots, l(k) - 1. \tag{59}
$$

From (57), (58), and Assumption 2, we have for all $k \in \mathcal{K}$

$$
\begin{aligned}
\frac{\varepsilon}{2} < \|\nabla f(x_1^k) - \nabla f(x_1^{l(k)})\| &\leq mL \|x_1^k - x_1^{l(k)}\| \\
&\leq mL \sum_{l=k}^{l(k)-1} \|x_1^{l+1} - x_1^l\| \\
&= mL \sum_{l=k}^{l(k)-1} \alpha^l \|D_m^l h_m^l\| \\
&\leq \frac{mL}{c_1} \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l} \|h_m^l\|,
\end{aligned}
$$

where the last inequality follows from (34). Thus, we obtain

$$\liminf_{k\to\infty, k\in\mathcal{K}} \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l}\|h_m^l\| > 0. \tag{60}$$

Let an infinite set $\mathcal{K}^+$ be defined by

$$\mathcal{K}^+ = \bigcup_{k\in\mathcal{K}} \{k, k+1, \ldots, l(k)-1\}.$$

Since $\|\nabla f(x_1^l)\| > \frac{\varepsilon}{2}, \forall l \in \mathcal{K}^+$ from (59), it follows that

$$\liminf_{l\to\infty, l\in\mathcal{K}^+} \|\nabla f(x_1^l)\| > 0. \tag{61}$$

If $\liminf_{l\to\infty, l\in\mathcal{K}^+} \|h_m^l\| = 0$, then from Lemma 7, we have $\liminf_{l\to\infty, l\in\mathcal{K}^+} \|\nabla f(x_1^l)\| = 0$. It contradicts (61). Therefore, there exists an $\varepsilon' > 0$ such that

$$\liminf_{l\to\infty, l\in\mathcal{K}^+} \|h_m^l\| > \varepsilon'. \tag{62}$$

For each $k \in \mathcal{K}$, we have from (52) that

$$f(x_1^{l(k)}) - f(x_1^k) \le -\sigma_1 \sum_{l=k}^{l(k)-1} \frac{\alpha^l\|h_m^l\|^2}{l} + \sigma_2 \sum_{l=k}^{l(k)-1} \frac{\alpha^l\|h_m^l\|}{(l)^2}.$$

It follows from (62) and Lemma 3 (III) that for any sufficient large $k \in \mathcal{K}$

$$\begin{aligned} f(x_1^{l(k)}) - f(x_1^k) &\le -\sigma_1 \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l}(\varepsilon')^2 + \sigma_2 \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{(l)^2}\rho_3 \\ &= -\sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l}\Big(\sigma_1(\varepsilon')^2 - \sigma_2\frac{\rho_3}{l}\Big). \end{aligned}$$

Since for a sufficient large $\bar{k}$, there exists an $\varepsilon'' > 0$ such that

$$\sigma_1(\varepsilon')^2 - \sigma_2\frac{\rho_3}{l} > \varepsilon'', \qquad \forall l > \bar{k},$$

we have that for a sufficient large $k \in \mathcal{K}$

$$f(x_1^{l(k)}) - f(x_1^k) \le -\varepsilon'' \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l}.$$

Since the sequence $\{f(x_1^k)\}$ converges from Lemma 8, if follows that

$$\lim_{k\to\infty, k\in\mathcal{K}} \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l} = 0.$$

From the boundedness of $\|h_m^k\|$, we have

$$\liminf_{k\to\infty, k\in\mathcal{K}} \sum_{l=k}^{l(k)-1} \frac{\alpha^l}{l}\|h_m^l\| = 0,$$

which contradicts (60). Therefore, there cannot exist an infinite set $\mathcal{K}$ such that $\|\nabla f(x_1^k)\| > \varepsilon$, $\forall k \in \mathcal{K}$, which means that $\lim_{k\to\infty} \|\nabla f(x_1^k)\| = 0$. **QED**

5 Implementation Issue of EKF-S

In the previous section, we have shown that EKF-S has the global convergence property under Assumptions 1-5. However a concrete update rule of α^k has yet to be specified. Note that the exact value of α_*^k in (29) can only be determined at the end of the kth iteration. Therefore we used to devise an implementable stepsize rule that still guarantees convergence of the algorithm.

Now we state the algorithm with an implementable stepsize rule. Later we will show that the stepsize rule is well-defined.

The Extended Kalman Filter with Adaptive Stepsize Rule (EKF-AS)

Step.0 Choose the initial point $x_1^1 \in \Re^n$, parameters $c > 0$, $\delta^1 > 0$ and $0 < \gamma < 1$. Set $k := 1$, $i := 1$, and $H_0^1 := \delta^1 I$. Choose $\alpha_*^0 \geq 1$.

Step.1 Choose λ^k such that $0 \leq 1 - (\lambda^k)^m \leq c/k$. Determine $\alpha^k(1)$ from α_*^{k-1} by

$$\alpha^k(1) := \max(1, \gamma \alpha_*^{k-1}). \tag{63}$$

Set $j := 1$, $\alpha_*^k(1) := \alpha_*^{k-1}$.

Step.2 Obtain x_{i+1}^k from

$$x_{i+1}^k := x_i^k - \alpha^k(j)(H_i^k)^{-1} \nabla g_i(x_i^k) g_i(x_i^k), \tag{64}$$

where

$$H_i^k := \lambda^k H_{i-1}^k + \nabla g_i(x_i^k) \nabla g_i(x_i^k)^T. \tag{65}$$

Step.3 If $i = m$, then go to **Step.4**. Otherwise, set $i := i + 1$ and go to **Step.2**.

Step.4 If a termination criterion is satisfied, then stop. Otherwise, go to **Step.5**.

Step.5 Set

$$\alpha_*^k(j) = \begin{cases} \frac{(1-\eta)}{L} \frac{(x_1^{k+1} - x_1^k)^T H_m^k (x_1^{k+1} - x_1^k)}{\|x_1^{k+1} - x_1^k\| \sum_{i=2}^m \|x_i^k - x_1^k\| + \frac{m}{2} \|x_1^{k+1} - x_1^k\|^2}, & \text{if} \quad x_1^{k+1} \neq x_1^k \\ 0, & \text{otherwise.} \end{cases}$$

If

$$1 \leq \alpha^k(j) \leq \max(1, \alpha_*^k(j)), \tag{66}$$

then set $\alpha_*^k = \alpha_*^k(j)$ and go to **Step.6**. Otherwise set $\alpha^k(j+1) := \max(1, \gamma \alpha_*^k(j))$, $j := j + 1$, $i := 1$ and go to **Step.2**.

Step.6 Choose $\delta^{k+1} \geq 0$ and set

$$x_1^{k+1} := x_{m+1}^k, \quad H_0^{k+1} := H_m^k + \delta^{k+1} I.$$

Set $i := 1$, $k := k + 1$, and go to **Step.1**.

The next theorem shows that Step.2–Step.5 in EKF-AS cannot be repeated infinitely for any k.

Theorem 2 *For each k, there exists a positive integer j^k such that* $1 \leq \alpha^k(j^k) \leq \max(1, \alpha_*^k(j^k))$.

Proof : Consider the kth iteration. If $\alpha_*^k(j) < 1$ for some j, then $\alpha^k(j) = 1$, and hence (66) holds. Now assume that for all j, $\alpha_*^k(j) \geq 1$ and (66) never holds. Then, from (66) and the updating rule of $\alpha^k(j)$, we have

$$1 \leq \alpha_*^k(j) < \alpha^k(j) = \gamma\alpha_*^k(j-1).$$

Thus, we obtain

$$\alpha_*^k(j) < (\gamma)^{j-1}\alpha_*^k(1).$$

Since we have $\alpha_*^k(j) < 1$ for a sufficient large j, this contradicts the above assumption. The proof is complete. **QED**

6 Numerical Experiments

In this section, to evaluate the practical performance of the EKF-AS proposed in the previous section, we implemented the algorithm. First, we compared the performance of EKF-AS with EKF proposed by Bertsekas [1]. Next, we investigated the behavior of stepsizes determined by EKF-AS.

6.1 Test Problems

We have solved test problems that arise in Multilayer Neural Network (MNN) training (See Appendix A in detail). More specifically, we used 3-layer neural network with 25 input units, 4 hidden units and 4 output units, and the sigmoid function $G : \Re \mapsto \Re$ defined by

$$G(\chi) = \frac{1}{1+\exp(-\chi)}.$$

We trained MNN for two pattern recognition problems described as follows.

Problem 1 : 25 bit alphabet recognition problem (no error data)

This problem is a learning problem of recognizing 10 alphabet characters. The input and the output patterns are given by

$$I_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad O_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ldots, I_{10} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}, \quad O_{10} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

where alphabets A ~ J are represented as 5×5 matrices $I_i, i = 1, \ldots, 10$ with 0-1 components, and $O_i, i = 1, \ldots, 10$ are binary representations for integers 0 ~ 9. The term "no error data" means that each input pattern always corresponds to the right output pattern. In this case, it is possible for MNN to be trained correctly, that is, $\min_x f(x) = 0$ is obtained.

Problem 2 : 25 bit alphabet recognition with error data

This problem is the almost same as Problem 1. Its input patterns $\tilde{I}_1, \ldots, \tilde{I}_{10}$ are given by

$$\tilde{I}_i = I_i, \quad i \neq 8, 10, \qquad \tilde{I}_8 = I_2, \quad \tilde{I}_{10} = I_5.$$

On the other hand, the output patterns $\tilde{O}_1, \ldots, \tilde{O}_{10}$ are the same as Problem 1, that is $\tilde{O}_i = O_i, i = 1, \ldots, 10$. Therefore it contains error data, and hence $\min_x f(x) > 0$.

6.2 Experimental Results

We chose the parameters that are common to EKF-AS and EKF as follows:

$$\delta^k = \begin{cases} 0.5, & \text{if} \quad k = 1 \\ 0.05, & \text{if} \quad k > 1 \end{cases}, \qquad \lambda^k = \max\left(1 - \frac{5}{k}, 0.8\right)$$

Furthermore we set the constants in EKF-AS as $\gamma = 0.95, \eta = 0.1, L = 0.018$. We employed the termination criterion:

$$\left\| \sum_{i=1}^{m} \nabla g_i(x_i^k) g_i(x_i^k) \right\| < \sum_{i=1}^{m} r_i \times 5.0 \times 10^{-6}, \tag{67}$$

where is also common to EKF and EKF-AS. All computer programs were coded in MATLAB Version 5.3.1 and were run on a SUN ULTRA 60. Each component of the starting point was randomly generated from the interval $[-1, 1]$. To measure the efficiency of the algorithms, we used two kinds of average. One is the usual average, which we call *standard* average, and the other is the average using reciprocal number, which we call *inverse* average. The inverse average p of $p_1, \ldots, p_n$ is defined by

$$p = \frac{1}{\left(\frac{1}{p_1} + \cdots + \frac{1}{p_n}\right)/n} = \frac{n}{\frac{1}{p_1} + \cdots + \frac{1}{p_n}}.$$

Table 1 gives the standard average and inverse average of the number of iterations over 10 runs, of the two methods on the two test problems. In Table 1, $(\cdot)$ denotes the average number of backtrackings to find α^k satisfying (66) in EKF-AS. Table 1 indicates that EKF-AS is consistently more effective than EKF. However, the efficiency of EKF-AS for Problem 2 is not so remarkable as that for Problem 1. The reason may be explained as follows. Since Problem 2 contains error data, $\sum_{i=1}^{m} \|x_i^k - x_1^k\|$ may not converge to 0, and hence stepsize remains relatively small. Therefore, EKF-AS behaves like EKF. Next we discuss the number of backtrackings to determine stepsize. Table 1 shows that the average number of backtrackings is less than 10, which is small relative to the total number of iterations.

Table 1: Results for two problems

	Algorithm	EKF-AS	EKF
Problem.1	standard average	68.6 (4.3)	137.7
	inverse average	67.2	133.6
	min – max	54 – 86	100 – 180
Problem.2	standard average	129.4 (6.3)	169.4
	inverse average	126.3	161.7
	min – max	95 – 163	134 – 239

Figures 1 and 2 show the behavior of stepsizes in EKF-AS. From Figures 1 and 2, we see that the stepsize is relatively large in the early stage of the iterations. The reason for this may be explained as follows. Since the iterates in the early stage are usually far from a stationary point of the problem, each component of $\nabla g_i(x)$ is generally large, and hence the norm of H_i^k is large. Then, since the numerator of (29) tends to be large, α^k become large. Moreover, Figures 1 and 2 show that the stepsize increases linearly at the later stage of the iterations.

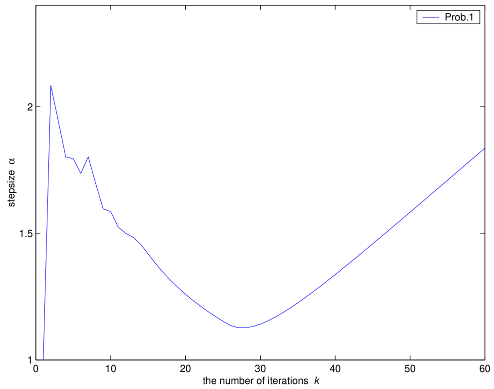


Figure 1: The behavior of α^k for Problem 1

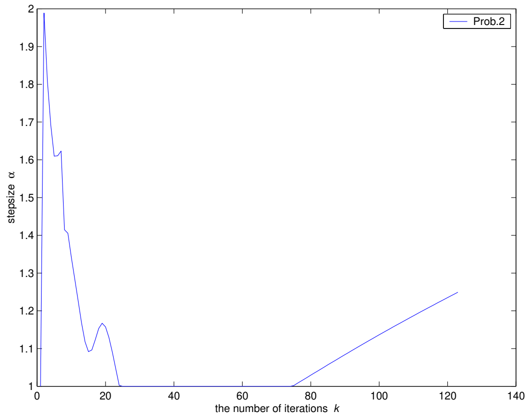


Figure 2: The behavior of α^k for Problem 2

7 Conclusion

In this paper, we have aimed to improve the convergence speed of EKF, and proposed the algorithm called EKF with stepsize. Moreover, we have proved that the proposed method has the global convergence property under appropriate assumptions. Finally, we have conducted numerical experiments with the proposed method to evaluate the practical performance. The numerical results indicate that our method is promising.

One of further subjects for research is the analysis of the convergence rate of the proposed method. Another important task is to develop efficient rules for updating parameters λ^k and δ^k, since the convergence speed of EKF-AS and EKF substantially depend on the choice of λ^k and δ^k.

References

[1] D. P. Bertsekas, *Incremental Least Squares Methods and the Extended Kalman Filter*, SIAM Journal on Optimizaiton, Vol. 3, pp. 807-822, 1996.

[2] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1996.

[3] D. P. Bertsekas, *A New Class of Incremental Gradient Methods for Least Squares Problems*, SIAM Journal on Optimization, Vol. 2, pp. 913-926, 1997.

[4] J. E. Dennis, Jr, R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[5] Z. Q. Luo and P. Tseng, *Analysis of an Approximate Gradient Projection Method with Applications to the Back Propagation Algorithm*, Optimization Methods and Software, Vol. 4, pp. 85-101, 1994.

[6] O. L. Mangasarian and M. V. Solodov, *Serial and Parallel Backpropagation Convergence via Nonmonotone Pertubed Minimization*, Optimization Methods and Software, Vol. 4, pp. 103-116, 1994.

[7] W. S. McCulloch and W. Pitts, *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bull. Math. Biophysics, Vol. 5, pp. 115-133, 1943.

[8] T. N. Pappas, *Solution of Nonlinear Equations by Davidon's Least Squares Method*, M.S. Thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 1982.

[9] B. T. Poljak, *Introduction to Optimization*, Optimization Software Inc., NY, 1987.

[10] D. E. Rumelhart, G.E. Hinton, and R.J. Williams, *Learning Internal Representations by Error Propagation*, Parallel Distributed Processing, MIT Press, pp. 318-362, 1986.

[11] P. Tseng, *Incremental Gradient(-Projection) Method with Momentum Term and Adaptive Stepsize Rule*, SIAM Journal on Optimizaiton, Vol. 8, pp. 506-531, 1998.

A Multilayer Neural Network

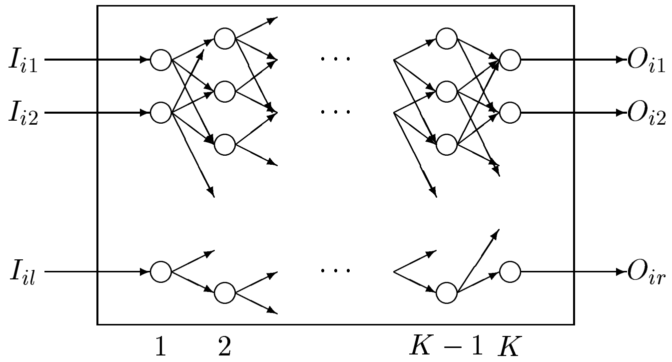
In the appendix, we describe the multilayer neural network (MNN) [7]. We consider a network with units in K layers (Figure 3). Regarding the first layer as the input layer, the ith unit on the kth layer is governed by the following equation

$$y_k^i = G\left(\sum_{j=1}^{n_{k-1}} u_{k-1}^{ji} y_{k-1}^j - \theta_k^i\right), \qquad k = 2, \dots, K, \ i = 1, \dots, n_k$$

where n_k is the number of units in the kth layer, θ_k^i is a threshold value of the ith unit in the kth layer, u_k^{ji} is a weight for the arc from the jth unit in the kth layer to the ith unit in the $k+1$th layer, and $G(\chi)$ is the sigmoid function given by

$$G(\chi) = \frac{1}{1 + \exp(-\gamma\chi)},$$

where γ is a positive constant. In the input layer, the output signal from a unit is equal to the input to the unit. When a vector I is given to the input layer, the output vector from the output layer is regarded as a function of weights $u = (u_1^{11}, \dots, u_K^{n_{K-1}n_K})^T$, threshold values $\theta = (\theta_1^1, \dots, \theta_K^{n_K})^T$ and the input vector I. Therefore, letting $x := (u^T, \theta^T)^T \in \Re^n$, we denote the output vector as $z(x, I)$.



$I_i = (I_{i1}, I_{i2}, \dots, I_{il})^T$, $z_i = (z_{i1}, z_{i2}, \dots, z_{ir})^T$

Figure 3: K-layer neural network

The learning of MNN means a process of adjusting the weights and threshold values $x = (u, \theta)$ to realize given m input-output pairs

$$\begin{array}{ccc} \text{input vectors} & & \text{target vectors} \\ I_1 & \longrightarrow & O_1 \\ I_2 & \longrightarrow & O_2 \\ \vdots & & \vdots \\ I_m & \longrightarrow & O_m \end{array}$$

where each input vector $I_p \in \Re^l$ is called a 'pattern' and the corresponding target vector $O_p \in \Re^r$ is called the 'training signal' for pattern p. To measure a gap between the 'training signal' O_i and the output $z_i(x) := z(x, I_i)$ of MNN, we define the output error function as

$$g_i(x) = z_i(x) - t_i \ , \qquad i = 1, \dots, m,$$

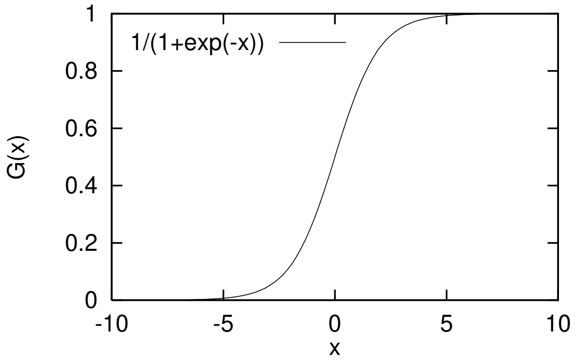


Figure 4: Sigmoid function

and denote $g(x) := (g_1(x), \ldots, g_m(x))$. Then, the total error function $f(x)$ for all inputs is given by

$$f(x) = \frac{1}{2} \|g(x)\|^2 = \sum_{i=1}^{m} \frac{1}{2} \|g_i(x)\|^2 .$$

Thus, the learning of MNN can be formulated as the minimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \Re^n. \end{array}$$