⟨Research Paper⟩

Journal of the Korean Statistical Society
Vol.12, No.1, 1983

# Nonlinear Regression with Censored Data

D.W. Shin* and D.S. Bai**

## ABSTRACT

An algorithm based on EM procedure which finds maximum likelihood estimators in a nonlinear regression with censored data is proposed, and asymptotic properties of the estimator are investigated in detail. Some numerical examples are also given.

## I. Introduction

In a recent paper, Aitken (1981) utilizes the EM procedure (Dempster et al. (1977)) for estimating parameters in linear models with right censored normally distributed response variables. In this paper, his results are extended to the case where the substantive nature of the data or other considerations indicate a response function which is not linear in the unknown parameters.

Consider a nonlinear regression model

$$y_i=f(x_i,\boldsymbol{\theta})+\eta_i,\quad i=1,\cdots,N, \tag{1}$$

where $f(x_i,\boldsymbol{\theta})$ is a known function except for a $k\times 1$ vector parameter $\boldsymbol{\theta}$, and $\eta_1,\cdots,\eta_N$ are i.i.d. normal random variables with

$$E(\eta_1)=0,\quad \mathrm{Var}(\eta_1)=\sigma^2.$$

For each $x_i$, $i=1,\cdots,N$, $y_i$ is observed if it is not greater than $y_i^*$, otherwise, it is censored at $y_i^*$, where $y_i^*$ is a predetermined constant. Let the number of censored data be $m$, $1\leq m\leq N$. Without loss of generality, we assume that the observations are rearranged as

* Daelim Industrial Co. Ltd.
** Korea Advanced Institute of Science and Technology

$$y_1, \cdots, y_n, \quad y_{n+1}^*, \cdots, y_N^*,$$

where $n=N-m$.

In Section Ⅱ, an iterative procedure is proposed for obtaining maximum likelihood estimators (MLE's) of $\boldsymbol{\theta}$ and $\sigma^2$ in (1) via EM algorithm. Consistency and asymptotic normality of the estimators are studied in Section Ⅲ. And numerical examples are given in Section Ⅳ.

## Ⅱ. Maximum Likelihood Estimation via EM Algorithm

The EM procedure proposed by Dempster et al.(1977) is useful for maximum likelihood estimation with incomplete data. The algorithm consists of two parts; expectation step and maximization step. In the expectation step, the conditional expectations of the incomplete data are obtained and in the maximization step the MLE's are obtained regarding the conditional expectations as complete data.

The following algorithm is an application of the EM procedure to obtain MLE's in (1).

### Algorithm

**Step 1** Give initial values, $\boldsymbol{\omega}^{(1)}=(\boldsymbol{\theta}^{(1)}, \sigma^{(1)})$, of $\boldsymbol{\omega}=(\boldsymbol{\theta}, \sigma)$. Set $\nu=1$.

**Step 2** For $i=n+1, \cdots, N$, compute

$E(Y_i \mid Y_i > y_i^*, \boldsymbol{\omega}^{(\nu)}) = f(x_i, \boldsymbol{\theta}^{(\nu)}) + \sigma^{(\nu)} S(z_{i\nu}^*)$,

where $S(t)=\phi(t)/\Psi(t)$,

$$\phi(t)=(2\pi)^{-1/2}\exp(-t^2/2),$$

$$\Psi(t)=\int_t^\infty \phi(u)du,$$

and

$$z_{i\nu}^*=\{y_i^*-f(x_i, \boldsymbol{\theta}^{(\nu)})\}/\sigma^{(\nu)}.$$

**Step 3** a) Obtain the least square estimators, $\boldsymbol{\theta}^{(\nu+1)}$, of $\boldsymbol{\theta}$ regarding $E(Y_i \mid Y_i > y_i^*, \boldsymbol{\omega}^{(\nu)})$ as complete data for $i=n+1, \cdots, N$.

b) Compute

$$\sigma^{(\nu+1)}=[\sum_1^n (y_i-f(x_i, \boldsymbol{\theta}^{(\nu)}))^2+\sigma^{(\nu)^2} \sum_{n+1}^N (1+u_i S(u_i))]^{1/2}/N^{1/2}, \quad (2)$$

where $u_i=\{y_i^*-f(x_i, \boldsymbol{\theta}^{(\nu+1)})\}/\sigma^{(\nu)}$.

Check convergence. If the convergence criterion is satisfied, stop. Otherwise, set $\nu=\nu+1$ and go to Step 2.

Let $L$ be the likelihood function, i.e.,

$$L=\left[\prod_1^n \frac{1}{\sigma}\phi(z_i)\right]\left[\prod_{n+1}^N \Psi(z_i^*)\right],$$

where $z_i=(y_i-f(x_i,\boldsymbol{\theta}))/\sigma,\ i=1,\cdots,n,$

and $z_i^*=(y_i^*-f(x_i,\boldsymbol{\theta}))/\sigma,\ i=n+1,\cdots,N.$

Differentiating $\log L$ with respect to $\boldsymbol{\omega}$, we obtain

$$\frac{\partial \log L}{\partial \theta_j}=\frac{1}{\sigma^2}\sum_1^N (w_i-f(x_i,\boldsymbol{\theta}))\frac{\partial f(x_i,\boldsymbol{\theta})}{\partial \theta_j}=0,\ j=1,\cdots,k, \qquad (3)$$

$$\frac{\partial \log L}{\partial \sigma}=\frac{1}{\sigma^3}\left(n-\sum_{n+1}^N z_i^* S(z_i^*)\right)\left[\frac{\sum_1^n (y_i-f(x_i,\boldsymbol{\theta}))^2}{n-\sum_{n+1}^N z_i^* S(z_i^*)}-\sigma^2\right]=0, \qquad (4)$$

where $w_i=\begin{cases} y_i, & i=1,\cdots,n, \\ f(x_i,\boldsymbol{\theta})+\sigma S(z_i^*), & i=n+1,\cdots,N. \end{cases}$

Note that (3) is just the normal equation of the nonlinear regression model (1) if $(w_{n+1},\cdots,w_N)$ is considered to be complete data. Also, if we let $\boldsymbol{\theta}^{(\nu+1)}=\boldsymbol{\theta}^{(\nu)}=\boldsymbol{\theta}$ and $\sigma^{(\nu+1)}=\sigma^{(\nu)}=\sigma$, (2) reduces to (4). Consequently, the algorithm is just an iterative procedure to solve the normal equations (3) and (4), which implies that the limit of the sequence of the estimates obtained by the above procedure is the MLE if it exists. The convergence of the EM algorithm was proved by Dempster et al. (1977) and Boyles (1983).

## III. Asymptotic Properties of the MLE's

Since the samples in (1) are censored, they are not necessarily identically distributed. Hence, the usual limit theorems for i.i.d. random variables are not applicable. Hoadley (1971) studied the asymptotic properties of MLE's in INID (Independent Not Identically Distributed) cases and derived the regularity conditions for consistency and asymptotic normality. Following him, the limiting behaviors of the proposed estimators are investigated in this section.

The following notations will be used throughout the section.

$\|\cdot\|$; Eucleadian norm,

$\boldsymbol{\omega}$; $(\boldsymbol{\theta},\sigma)$, $(k+1)\times 1$ vector parameter,

$\boldsymbol{\omega}^0$; $(\boldsymbol{\theta}^0, \sigma^0)$, true value of $\boldsymbol{\omega}$,

$\hat{\boldsymbol{\omega}}_N$; MLE of $\boldsymbol{\omega}$,

$\Theta$; closed subset of $R^k$,

$\Omega$; $\{\boldsymbol{\omega} : \boldsymbol{\theta} \in \Theta,\ \sigma \geq \varepsilon\}$, where $\varepsilon > 0$ and known,

$g_i(y_i|\boldsymbol{\omega})$; p.d.f. of $Y_i$, $i=1, \cdots, N$,

$$R_i(\boldsymbol{\omega}) = \begin{cases} \log\{g_i(y_i|\boldsymbol{\omega})/g_i(y_i|\boldsymbol{\omega}^0)\}, & \text{when } g_i(y_i|\boldsymbol{\omega}^0) > 0, \\ 0 & \text{, when } g_i(y_i|\boldsymbol{\omega}^0) = 0, \end{cases}$$

$R_i(\boldsymbol{\omega}, \rho) = \sup\{R_i(\boldsymbol{t}) : ||\boldsymbol{t} - \boldsymbol{\omega}|| < \rho\}$, where $\rho$ is a positive constant,

$V_i(r) = \sup\{R_i(\boldsymbol{\omega}) : ||\boldsymbol{\omega}|| > r\}$, where $r$ is a positive constant,

$$Y^{(B)} = \begin{cases} Y & \text{, when } Y \geq -B, \\ -B, & \text{when } Y < -B, \end{cases}$$

where $B$ is a nonnegative constant.

Also, for $Y_i^{(B)}$, let $R_i^{(B)}(\boldsymbol{\omega})$, $R_i^{(B)}(\boldsymbol{\omega}, \rho)$ and $V_i^{(B)}(r)$ be similarly defined.

In (1), the p.d.f of $Y_i$ may be rewritten as

$$g_i(y_i) = \begin{cases} \phi(z_i)/\sigma, & \text{when } y_i < y_i^*, \\ \Psi(z_i^*), & \text{when } y_i = y_i^*, \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where $z_i = (y_i - f(x_i, \boldsymbol{\theta}))/\sigma$, and $z_i^* = (y_i^* - f(x_i, \boldsymbol{\theta}))/\sigma$.

We now examine some lemmas useful for checking the consistency conditions of Hoadley (1971).

**Lemma 1.** For $\boldsymbol{\omega} \in \Omega$, there exist positive numbers $\rho^*$, $r$ and $M$ such that

$$E_{\omega^0}[R_i^{(0)}(\boldsymbol{\omega}, \rho)]^3 \leq M, \text{ for } 0,\ \rho \leqslant \rho^*,\ i=1, \cdots, N,$$

and

$$E_{\omega^0}[V_i^{(0)}(r)]^3 \leq M.$$

**Proof.** From (5), the pdf of $Y_i^{(0)}$ is given by

$$g_i^{(0)}(y_i^{(0)}|\boldsymbol{\omega}) = \begin{cases} \Phi(-f(x_i, \boldsymbol{\theta})/\sigma, & \text{for } y_i^{(0)} = 0, \\ \phi((y_i^{(0)} - f(x_i, \boldsymbol{\theta}))/\sigma), & \text{for } 0 < y_i^{(0)} < y_i^*, \\ \Psi((y_i^* - f(x_i, \boldsymbol{\theta}))/\sigma), & \text{for } y_i^{(0)} = y_i^*, \\ 0, & \text{otherwise,} \end{cases}$$

where $\Phi(\cdot) = 1 - \Psi(\cdot)$.

Hence, for all $\boldsymbol{\omega} \in \Omega$

$$\log[g_i^{(0)}(y_i^{(0)}|\boldsymbol{\omega})/g_i^{(0)}(y_i^{(0)}|\boldsymbol{\omega}^0)]$$

$$\leq \begin{cases} -\log\Phi(-f(x_i,\boldsymbol{\theta}^0)/\sigma^0), & \text{for } y_i^{(0)}=0, \\ (y_i^{(0)}-f(x_i,\boldsymbol{\theta}^0))/2(\sigma^0)^2-\log\varepsilon+\log\sigma^0, & \text{for } 0<y_i^{(0)}<y_i^*, \\ -\log\Psi(y_i^*-f(x_i,\boldsymbol{\theta}^0)/\sigma^0), & \text{for } y_i^{(0)}=y_i^*. \end{cases}$$

The right hand side is free of $\boldsymbol{\omega}$, which implies that there exists $M$ such that

$$E_{\omega^0}[R_i^{(0)}(\boldsymbol{\omega},\rho)]^3\leq M,$$

and

$$E_{\omega^0}[V_i^{(0)}(r)]^3\leq M.$$

**Lemma 2.** If the distribution of $Y_i$ under $\boldsymbol{\omega}\neq\boldsymbol{\omega}^0$ differs from that under $\boldsymbol{\omega}^0$, then

$$\lim_{N\to\infty}\sum_1^N E(R_i(\boldsymbol{\omega}))/N<0.$$

**Proof.** See Wald (1949).

**Lemma 3.** If $\lim_{||\boldsymbol{\theta}||\to\infty} f(x_i,\boldsymbol{\theta})=\infty$ for all $x_i$, $i=1,\cdots,N$,

$$\lim_{N\to\infty}\sum_1^N E(V_i(r))/N<0.$$

**Proof.** $\log[g_i(y_i|\boldsymbol{\omega})/g_i(y_i|\boldsymbol{\omega}^0)]$

$$\leq \begin{cases} -\frac{1}{2}(f(x_i,\boldsymbol{\theta})/\sigma)^2+(y_i^*)^2+\frac{1}{2}\left(\frac{y_i-f(x_i,\boldsymbol{\theta}^0)}{\sigma^0}\right)^2-\log\frac{\sigma}{\sigma^0}, & \text{for } y_i<y_i^*, \\ -\log\Psi((y_i^*-f(x_i,\boldsymbol{\theta}^0))/\sigma^0), & \text{for } y_i=y_i^*. \end{cases}$$

Therefore, $E_{\omega^0}\{\sup[\log\{g_i(y_i|\boldsymbol{\omega})/g_i(y_i|\boldsymbol{\omega}^0)\} : ||\boldsymbol{\omega}||>r]\}$ has an upper bound of the form

$$a \sup\ \{-f(x,\boldsymbol{\theta})^2/2\sigma^2-\log\sigma : ||\boldsymbol{\omega}||>r\}+b,\ \ a>0,$$

which goes to $-\infty$ as $||\boldsymbol{\omega}||\to\infty$

**Lemma 4.** $R_i(\boldsymbol{\omega},\rho)$ and $V_i(r)$ are measurable functions of $y_i$ for each $i$.

**Proof.** The measurability of $R_i(\boldsymbol{\omega},\rho)$ and $V_i(r)$ follows immediately from the measurability of $R_i(\boldsymbol{\omega})$ and its supremum.

The following theorem shows the consistency of the proposed estimators.

**Theorem 1.** If, for $\boldsymbol{\omega}\in\Omega$, $f(x_i,\boldsymbol{\theta})$ is a continuous function of $\boldsymbol{\theta}$ and the conditions of Lemmas 2 and 3 hold,

$$\hat{\boldsymbol{\omega}}_N \xrightarrow{p} \boldsymbol{\omega}^0$$

**Proof.** From Lemmas 1 through 4, it is easy to show that the conditions of Hoadley (1971) for consistency are satisfied.

In practice, the conditions of $\lim_{||\theta|| \to \infty} f(x, \boldsymbol{\theta}) = \infty$ and $\sigma \geq \varepsilon$ cause little difficulty since we are not intersted in the values of $\boldsymbol{\theta}$ sufficiently far from the origin in the parameter space. Moreover, in many cases, $\sigma \geq \varepsilon$ for a sufficiently small value of $\varepsilon > 0$.

We now introduce additional lemmas useful for checking the asymptotic normality conditions of Hoadley (1971).

**Lemma 5.** If $f(x_i, \boldsymbol{\theta})$ has a second derivative with respect to $\boldsymbol{\theta}$, then $\int_{-\infty}^{\infty} g_i(y_i|\boldsymbol{\omega})dy_i$ is twice differentiable under the integral sign.

**Proof.** Since

$$\int_{-\infty}^{\infty} g_i(y_i|\boldsymbol{\omega})dy_i = \int_{-\infty}^{y_i^*} g_i(y_i|\boldsymbol{\omega})dy_i + g_i(y_i^*|\boldsymbol{\omega})\Psi((y_i - f(x_i, \boldsymbol{\theta}))/\sigma),$$

it suffices to show that

$$\frac{\partial^2}{\partial\theta_j\partial\theta_l}\int_{-\infty}^{y_i^*} g_i(y_i|\boldsymbol{\omega})dy_i = \int_{-\infty}^{y_i^*} \frac{\partial^2}{\partial\theta_j\partial\theta_l} g_i(y_i|\boldsymbol{\omega})dy_i,$$

which follows from the fact that $g_i(y_i|\boldsymbol{\omega})$ belongs to an exponential family.

**Lemma 6.** There exists $M > 0$ such that

$$E_{\omega^\circ}[|\partial \log g_i(y_i|\boldsymbol{\omega})/\partial\theta_j|^3] \leq M, \quad j = 1, \cdots, k+1.$$

**Proof.** See the proof of Lemma 1.

**Lemma 7.** Under the condition of Lemma 5, there exist positive constants $\varepsilon', \delta$, and $M'$ and random variables $B_{i,\alpha\beta}(Y_i)$ such that

$$\sup_t \left\{\left|\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta} \log g_i(y_i|\boldsymbol{t})\right| : ||\boldsymbol{t} - \boldsymbol{\omega}|| < \varepsilon', \ \boldsymbol{t} \in \Omega\right\} \leq B_{i,\alpha\beta}(y_i)$$

and

$$E\{|B_{i,\alpha\beta}(Y_i)|^{1+\delta}\} \leq M',$$

for $i = 1, \cdots, N$, and $\alpha, \beta = 1, \cdots, k+1$.

**Proof** Define $A_{i,\alpha\beta}$, $C_{i,\alpha\beta}$ and $D_{i,\alpha\beta}$ as follows.

$$A_{i,\alpha\beta} = \sup\{|(f_\alpha(x_i, \boldsymbol{t})f_\beta(x_i, \boldsymbol{t}) - f(x_i, \boldsymbol{t})f_{\alpha\beta}(x_i, \boldsymbol{t}))/2\, t_{k+1}^2| : \boldsymbol{t} \in T\},$$

$$C_{i,\alpha\beta} = \sup\{|(f(x_i, \boldsymbol{t})f_{\alpha\beta}(x_i, \boldsymbol{t}))/2\, t_{k+1}^2| : \boldsymbol{t} \in T\},$$

and

$$D_{i,\alpha\beta} = \sup\{|S'(z_i^*)f_\alpha(x_i, \boldsymbol{t}) - S(z_i^*)f_{\alpha\beta}(x_i, \boldsymbol{t})| : \boldsymbol{t} \in T\},$$

for $i = 1, \cdots, N$, and $\alpha, \beta = 1, \cdots, k+1$,

where $f_\alpha = \frac{\partial}{\partial\theta_\alpha} f$, $f_{\alpha\beta} = \frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta} f$, $S'(t) = dS(t)/dt$,

and the supremum is taken over the set $T = \{\boldsymbol{t} : ||\boldsymbol{t} - \boldsymbol{\omega}|| \leq \varepsilon', \ \boldsymbol{t} \in \Omega\}$.

$$\text{Let } B_{i,\alpha\beta}(y_i)=\begin{cases} A_{i,\alpha\beta}+C_{i,\alpha\beta}\, y_i & \text{for } y_i < y_i^*, \\ D_{i,\alpha\beta}, & \text{for } y_i=y_i^*. \end{cases}$$

Clearly,

$$\left|\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta}\log g_i(y_i|\boldsymbol{\omega})\right|\leq B_{i,\alpha\beta}(y_i)$$

and $E|B_{i,\alpha\beta}(y_i)|^{1+\delta}\leq M'$ for some $\delta>0$ and $M'>0$.

**Lemma 8.** If the condition of Lemma 5 is satisfied, then $\log g_i(y_i|\boldsymbol{\omega})$ has a continuous and measurable second derivative with respect to $\boldsymbol{\omega}$, $i=1,\cdots,N$.

**Proof.** Obvious

The following theorem deals with the asymptotic normality of the proposed estimators.

**Theorem 2.** Let

$$\Gamma_N(\boldsymbol{\omega})=1/N\sum_1^N E_\omega\left[\frac{\partial}{\partial\boldsymbol{\omega}}\log g_i(Y_i|\boldsymbol{\omega})\left(\frac{\partial}{\partial\boldsymbol{\omega}}\log g_i(Y_i|\boldsymbol{\omega})\right)^T\right]$$

and $\Gamma(\boldsymbol{\omega})=\lim_{N\to\infty}\Gamma_N(\boldsymbol{\omega})$. In (1), if $\boldsymbol{\omega}_N\xrightarrow{p}\boldsymbol{\omega}^0$ and $\Gamma(\boldsymbol{\omega})$ is positive definite,

$$\sqrt{N}(\hat{\boldsymbol{\omega}}_N-\boldsymbol{\omega}^0)\xrightarrow{d}N(0,\Gamma^{-1}(\boldsymbol{\omega}^0))$$

under the condition of Lemma 5.

**Proof.** From Lemmas 5 through 8, it can be shown that the conditions of Hoadley (1971) for asymptotic normality are satisfied.

**Remark.** To test the hypothesis

$$H_0: \theta_1=\theta_1^0,\cdots,\ \theta_s=\theta_s^0,$$

the likelihood ratio

$$\lambda=\sup_{\Theta^*}\prod_1^N g_i(y_i|\boldsymbol{\omega})/\sup_{\Omega}\prod_1^N g_i(y_i|\boldsymbol{\omega})$$

can be used as the test statistic, where $\Theta^*=\{\boldsymbol{\omega}:\boldsymbol{\omega}\in\Omega,\ \theta_1=\theta_1^0,\cdots,\theta_s=\theta_s^0\}$. If the conditions of Theorem 2 holds, then

$$-2\log\lambda\xrightarrow{d}\chi^2(s).$$

## Ⅳ. Numerical Examples

In this section, the proposed algorithm is studied with two examples. In order to use the algorithm, however, another algorithm is needed to obtain least squares estimates in the maximization step. Many such algorithms are available; Hartley's(1961) method, steepest decent method, Marquardt's (1963) method, etc. In this study, Marquardt's method is adopted since it is most frequently used.

**Example 1.** Table 1 gives the results of temperature accelerated life tests on electrical insulation in 40 motoretts, originally reported by Crawford(1970). Ten motorettes were tested at each of four temperatures. Testing was terminated at different times at each temperature, resulting in a total of 17 failed units and 23 unfailed ones. The model used to analyze the data assumes that;

i) for any temperature, the distribution of time to failure is lognormal,

ii) the standard deviation, $\sigma$, of the time to failure is constant, and

iii) the mean, $\mu_x$, of the logarithm of the time to failure is a nonlinear function of the following form

$$\mu_x=\theta_1+\theta_2 x+\exp(\theta_3 x)-1,$$

where $x=1000/(T+273.2)$, $T$ is the absolute absolute temperature, and $\theta_1, \theta_2, \theta_3$ and $\sigma$ are unknown parameters.

**TABLE 1. Insulation life in hours at various test temperatures**

| 150° C | 170° C | 190° C | 220° C |
|---|---|---|---|
| | 1764 | 408 | 408 |
| | 2772 | 408 | 408 |
| | 3444 | 1344 | 504 |
| | 3542 | 1344 | 504 |
| | 3780 | 1440 | 504 |
| | 4860 | | |
| | 5196 | | |

1) All 10 motoretts at 150°C still on test without failure at 8064 hours.

2) 3 motoretts at 170°C still on test without failure at 5448 hours.

3) 5 motoretts at 190°C still on test without failure at 1680 hours.

4) 5 motoretts at 220°C still on test without failure at 528 hours.

Table 2 shows the results of the proposed method for obtaining MLE's of $\theta_1$, $\theta_2$, $\theta_3$ and $\sigma$. The final values are $\hat{\theta}_1=-5.799$, $\hat{\theta}_2=4.129$, $\hat{\theta}_3=0.076$ and $\hat{\sigma}=0.265$.

Table 3 shows the values of each parameters when the model is assumed to be $\mu_x=\theta_1+\theta_2 x$ as in Schmee and Hahn(1979). The final values are $\hat{\theta}_1=-6.036$, $\hat{\theta}_2=4.321$ and

$\hat{\sigma}=0.268$.

Comparing these values and in view of the fact that the data used is linear in nature, the proposed method seems to perform well.

**TABLE 2.** **MLE's of parametrs**
$(\mu_x=\theta_1+\theta_2 x+\{\exp(\theta_3 x)-1\})$

| Iteration | $\hat{\theta}_1$ | $\hat{\theta}_2$ | $\hat{\theta}_3$ | $\hat{\sigma}$ |
|---|---|---|---|---|
| 1 | −4.6008 | 4.0037 | −1.0241 | 0.1746 |
| 2 | −5.3222 | 3.8905 | 0.0635 | 0.2077 |
| 3 | −5.5140 | 3.9889 | 0.0660 | 0.2277 |
| 4 | −5.6322 | 4.0468 | 0.0693 | 0.2405 |
| 5 | −5.6826 | 4.0735 | 0.0706 | 0.2489 |
| 6 | −5.7107 | 4.0870 | 0.0728 | 0.2545 |
| 7 | −5.7401 | 4.0981 | 0.0767 | 0.2582 |
| 8 | −5.7687 | 4.1095 | 0.0787 | 0.2607 |
| 9 | −5.7687 | 4.1136 | 0.0759 | 0.2624 |
| 10 | −5.7690 | 4.1172 | 0.0736 | 0.2635 |
| 11 | −5.7991 | 4.1285 | 0.0755 | 0.2644 |
| 12 | −5.7992 | 4.1286 | 0.0760 | 0.2649 |

**TABLE 3.** **MLE's of parameters** (Schmee and Hahn (1969))
$(\mu_x=\theta_1+\theta_2 x)$

| Iteration | $\hat{\theta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}$ |
|---|---|---|---|
| 1 | −1.0000 | −1.0000 | 0.1748 |
| 2 | −4.9305 | 3.7470 | 0.2078 |
| 3 | −5.2871 | 3.9420 | 0.2277 |
| 4 | −5.5440 | 4.0733 | 0.2406 |
| 5 | −5.7125 | 4.1582 | 0.2493 |
| 6 | −5.8227 | 4.2135 | 0.2552 |
| 7 | −5.8957 | 4.2502 | 0.2593 |
| 8 | −5.9449 | 4.2749 | 0.2621 |
| 9 | −5.9784 | 4.2917 | 0.2641 |
| 10 | −6.0014 | 4.3033 | 0.2654 |
| 11 | −6.0173 | 4.3113 | 0.2664 |
| 12 | −6.0284 | 4.3168 | 0.2671 |
| 13 | −6.0361 | 4.3207 | 0.2676 |

**Example 2.** This example deals with some simulation results. The model and the experimental conditions studied are assumed to be the same as those of Example 1. At some points of the true parameter, $\boldsymbol{\omega}=(\theta_1, \theta_2, \theta_3, \sigma)$, 10 random numbers are generated

from the model for each temperature. The censoring time is assumed to be the expected time to failure at the temperature.

The MLE's of $\theta_1, \theta_2, \theta_3$ and $\sigma$ are computed via the proposed method. These are repeated 50 times for each $\boldsymbol{\omega}$ and the results are summarized in Table 4.

**TABLE 4. Mean and standard deviation of the estimates based on 50 simulation results**

| | $\hat{\theta}_1$ | $\hat{\theta}_2$ | $\hat{\theta}_3$ | $\hat{\sigma}$ |
|---|---|---|---|---|
| True value | −2.0 | 3.0 | −1.0 | 0.2 |
| Mean of MLE | −2.1365 | 3.0074 | −1.4575 | 0.1925 |
| Std. Dev. | 0.5601 | 0.2892 | 1.0202 | 0.0361 |
| True Value | −2.0 | 3.0 | −1.0 | 0.1 |
| Mean of MLE | −1.9649 | 2.9956 | −1.2951 | 0.0982 |
| Std. Dev. | 0.2907 | 0.1458 | 0.4756 | 0.0163 |
| True Value | 2.0 | 3.0 | −1.0 | 0.2 |
| Mean of MLE | 2.0149 | 3.0027 | −1.3258 | 0.1936 |
| Std. Dev. | 0.5923 | 0.2813 | 0.5649 | 0.0361 |
| True Value | 2.0 | 3.0 | −1.0 | 0.1 |
| Mean of MLE | 2.0232 | 3.0002 | −1.2138 | 0.0981 |
| Std. Dev. | 0.3047 | 0.1511 | 0.4916 | 0.0160 |
| True Value | −2.0 | 3.0 | 1.0 | 0.2 |
| Mean of MLE | −2.0882 | 3.0252 | 1.0010 | 0.1813 |
| Std. Dev. | 0.5242 | 0.2537 | 0.1156 | 0.0288 |
| True Value | −2.0 | 3.0 | 1.0 | 0.1 |
| Mean of MLE | −2.3413 | 3.0068 | 1.0017 | 0.0977 |
| Std. Dev. | 0.2789 | 0.1310 | 0.0059 | 0.0123 |
| True Value | −5.0 | −2.0 | 0.5 | 0.2 |
| Mean of MLE | −5.6410 | −1.4288 | 0.3907 | 0.1800 |
| Std. Dev. | 0.6748 | 0.3116 | 0.0582 | 0.0296 |
| True Value | −5.0 | −2.0 | 0.5 | 0.1 |
| Mean of MLE | −5.6135 | −1.4437 | 0.3769 | 0.0974 |
| Std. Dev. | 0.3533 | 0.1542 | 0.0192 | 0.0116 |

The table indicates that the algorithm performs fairly well. The estimated are somewhat sensitive to the initial values of the parameters. This, however, is an inherent problem in nonlinear regression situtions. It is partly due to the existence of $S(z_{i\nu}^*) = \phi(z_{i\nu}^*)/\Psi(z_{i\nu}^*)$ in Step 2. When initial values of the parameters differ from the true values considerably, $z_{i\nu}^* = (y_i - f(x_i, \boldsymbol{\theta}))/\sigma$ is so large that $\Psi(z_{i\nu}^*)$ becomes near 0 and the algorithm cannot go further. To prevent this, a search over several different initial values of the parameters is recommended.

## REFERENCES


1. Aitkin, M. (1981) A note on the Regression of Censored Data, *Technometrics*, Vol. 23, 161—163.
2. Boyles, R.A. (1983) On the Convergence of the EM algorithm, *J. Royal Stat. Soc. B*, Vol. 45, 47—50.
3. Crawford, D.E. (1970) Analysis of Life Test Data on Motorettes, *Insulation/ Circuit*, Vol. 16, No. 10, 43—48.
4. Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum Likelihood via the EM Algorithm, *J. Royal Stat. Soc. B*, Vol. 39, 1—23.
5. Hartley, H.O. (1961) Modified Gauss-Newton Method for Fitting of Nonlinear Regression Function, *Technometrics*, Vol. 3, 269—280.
6. Hoadley, B. (1969) Asymptotic Properties of Maximum Likelihood Estimators for the Independent Not Identically Distributed Case, *Ann. Math. Statist.*, Vol. 40, 633—643.
7. Marquardt, D.W. (1963) An Algorithm for Least Squares Estimation of Nonlinear Parameters, *J. Soc. for Industrial and Applied Math.*, Vol. 11, 431—441.
8. Schmee, J. and Hahn, C.J. (1979) A Simple Method for Regression Analysis with Censored Data, *Technometrics*, Vol. 21, 417—432.
9. Wald, A. (1949) Note on the Consistency of the Maximum Likelihood Estimate, *Ann. Math. Statist.*, Vol. 20, 595—601.