QUALITY ENGINEERING
2019, VOL. 31, NO. 3, 505–510
https://doi.org/10.1080/08982112.2018.1497181

Taylor & Francis
Taylor & Francis Group

Check for updates

OPEN CHALLENGES

An EM-algorithm approach to open challenges on correlation of intermediate and final measurements

Sumin Shen, Huiying Mao, and Xinwei Deng

Department of Statistics, Virginia Tech, Blacksburg, Virginia

Problem restatement

Product testing is a key element in the quality control. To ensure the final product quality, taking the intermediate measurements for identifying defects is an important step. However, when the measurements of products are destructive, it is difficult to quantify the correlation between the intermediate measurements and the final measurements because it is not possible to test the same product twice (Jensen 2018).

Let us denote the intermediate measurement as a random variable X and the final measurement as a random variable Y. Assume that the available data set contains m samples measured at the intermediate stage, $\boldsymbol{x}_{obs} = (x_1, x_2, \ldots, x_m)$, and $n - m$ samples measured at the final stage, $\boldsymbol{y}_{obs} = (y_{m+1}, y_{m+2}, \ldots, y_n)$. Suppose that a lower and upper limit specification of the final product is given by $\{L, U\}$. The problem of interest is to find the tolerance of intermediate measurements, $[a, b]$, to ensure that the specification will be met on the final measurements.

Our Solution: We address the aforementioned problem under the missing data framework. For the observed intermediate measurements, we denote their corresponding missing final measurements as $\boldsymbol{y}_{mis} = (y_1^*, \ldots, y_m^*)$. For the observed final measurements, we denote their corresponding missing intermediate measurements as $\boldsymbol{x}_{mis} = (x_{m+1}^*, \ldots, x_n^*)$. Then the paired complete data set $(\boldsymbol{x}, \boldsymbol{y})$ can be expressed as

$$\boldsymbol{x} = (\boldsymbol{x}_{obs}, \boldsymbol{x}_{mis}) = \left(x_1, \ldots, x_m, x_{m+1}^*, \ldots, x_n^*\right),$$
$$\boldsymbol{y} = (\boldsymbol{y}_{mis}, \boldsymbol{y}_{obs}) = \left(y_1^*, \ldots, y_m^*, y_{m+1}, \ldots, y_n\right).$$

We develop an estimation procedure using the expectation-maximization (EM) algorithm to quantify the relationship between the intermediate measurements and the final measurements. Consequently, we obtain the tolerance of the intermediate measurements given the specification on the final measurements. The proposed method can provide an accurate estimation and is robust to the initial values of the EM algorithm. The merits of the proposed method are illustrated by both simulation and the real-data case study.

The proposed method

We consider a linear regression between the intermediate measurements X and the final measurements Y, that is,

$$Y = \beta_0 + \beta_1 X + \epsilon,$$

where $\beta_1 \neq 0$ and the error term $\epsilon \sim N(0, \sigma^2)$. Furthermore, since the intermediate measurement is continuous, we assume that $X \sim N(\alpha, \delta^2)$ and X is independent of ϵ. Then the conditional distribution of Y given $X = x$ is

$$Y|X = x \sim N\left(\beta_0 + \beta_1 x, \sigma^2\right).$$

The joint distribution of X and Y is a bivariate normal distribution whose density function is given by

$$f_{XY}(x, y) = \quad f_{Y|X}\left(y|x\right) \cdot f_X(x).$$

Then the conditional distribution of X given $Y = y$ also follows a normal distribution $X|Y = y \sim N(E(X|Y = y), Var(X|Y = y))$ with

$$E\left(X|Y = y\right) = \alpha + \frac{\beta_1 \delta^2}{\sigma^2 + \beta_1^2 \delta^2}(y - \beta_0 - \beta_1 \alpha),$$
$$Var\left(X|Y = y\right) = \frac{\sigma^2 \delta^2}{\sigma^2 + \beta_1^2 \delta^2}.$$

Note that correlation coefficient between Y and X is closely connected with the slope β_1 in the simple

CONTACT Xinwei Deng xdeng@vt.edu Department of Statistics, Virginia Tech, Blacksburg, VA 24061, USA.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lqen.

linear regression. For a given value of $y \in [L, U]$, we can obtain the corresponding tolerance of X, $[a_y, b_y]$ such that

$$Pr\left(X \in \left[a_y, b_y\right] | y \in [L, U]\right) = 1 - \tau,$$

where the significant level τ is specified as 1 % in this study. Then we obtain the tolerance of X, $[a, b]$, by

$$a = min_y\{a_y\}, \quad b = max_y\{b_y\}.$$

As a result, we have

$$Pr(X \in [a, b] | y \in [L, U]) \geq Pr\left(X \in \left[a_y, b_y\right] | y \in [L, U]\right) = 1 - \tau.$$

The estimation of a and b need the estimation of parameters $\boldsymbol{\theta} = (\beta_0, \beta_1, \sigma^2, \alpha, \delta^2)$, which will be estimated by using the EM algorithm.

Expectation-maximization algorithm

The EM algorithm is commonly used to deal with missing data in various applications (Miljkovic and Barabanov, 2015). In the EM algorithm (Dempster, Laird, & Rubin 1977; Horton and Laird 1999), the E-step is to compute the expectation of the complete log-likelihood function through imputing the missing values by their conditional expectation, and the M-step is to estimate parameters by maximizing the expected complete log-likelihood function.

Given the full data $\boldsymbol{x} = (x_1, \ldots, x_m, x^*_{m+1}, \ldots, x^*_n)$ and $\boldsymbol{y} = (y^*_1, \ldots, y^*_m, y_{m+1}, \ldots, y_n)$, the complete log-likelihood function can be written as

$$l_c(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y}) = -n\log(2\pi) - n\log(\delta\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{m}\left(y^*_i - \beta_0 - \beta_1 x_i\right)^2 - \frac{1}{2\sigma^2}\sum_{i=m+1}^{n}\left(y_i - \beta_0 - \beta_1 x^*_i\right)^2 - \frac{1}{2\delta^2}\sum_{i=1}^{m}(x_i - \alpha)^2 - \frac{1}{2\delta^2}\sum_{i=m+1}^{n}\left(x^*_i - \alpha\right)^2.$$

In the **E-step**, we replace the missing values in $\boldsymbol{x}$ and $\boldsymbol{y}$ with the respective conditional expectation given the observed data, E $(x_i|y_i; \boldsymbol{\theta})$ and $E(y_i|x_i; \boldsymbol{\theta})$. Specifically,

$$Q(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y}) = E[l_c(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y})] = -n\log(2\pi) - n\log(\delta\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{m}\left[E\left(y_i^2|x_i; \boldsymbol{\theta}\right) - 2(\beta_0 + \beta_1 x_i)E\left(y_i|x_i; \boldsymbol{\theta}\right) + (\beta_0 + \beta_1 x_i)^2\right] - \frac{1}{2\sigma^2}\sum_{i=m+1}^{n}\left[(y_i - \beta_0)^2 + \beta_1^2 E\left(x_i^2|y_i; \boldsymbol{\theta}\right) - 2\beta_1(y_i - \beta_0)E\left(x_i|y_i; \boldsymbol{\theta}\right)\right] - \frac{1}{2\delta^2}\sum_{i=1}^{m}(x_i - \alpha)^2 - \frac{1}{2\delta^2}\sum_{i=m+1}^{n}\left[E\left(x_i^2|y_i; \boldsymbol{\theta}\right) - 2\alpha E\left(x_i|y_i; \boldsymbol{\theta}\right) + \alpha^2\right],$$

Table 1. Estimated parameters and tolerances under different initial slope in the real data.

$\beta_{1,init}$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\sigma}^2$	$\hat{\alpha}$	$\hat{\delta}^2$	[a, b]
−10.000	−0.27	41.22	0.60	41.31	28.17	[20.7, 62.5]
−6.667	−0.28	41.70	0.41	41.18	28.18	[21.0, 61.7]
−4.445	−0.29	42.08	0.28	41.13	28.09	[21.6, 61.0]
−2.223	−0.29	42.06	0.29	41.14	28.08	[21.5, 61.0]
−1.112	−0.29	42.05	0.30	41.17	28.04	[21.5, 61.1]
−0.001	−0.30	42.23	0.22	41.08	28.05	[21.8, 60.5]
0.001	0.30	18.26	0.22	41.08	28.05	[20.3, 59.0]
1.112	0.29	18.44	0.30	41.17	28.04	[20.0, 59.6]
5.556	0.27	19.10	0.53	41.26	28.19	[19.4, 60.8]

where the conditional expectation is given by

$$E(y_i|x_i) = \beta_0 + \beta_1 x_i, \; i = 1, \ldots, m.$$
$$E\left(y_i^2|x_i\right) = \sigma^2 + (\beta_0 + \beta_1 x_i)^2, \; i = 1, \ldots, m.$$
$$E\left(x_i^2|y_i\right) = \frac{\sigma^2\delta^2}{\sigma^2 + \beta_1^2\delta^2} + \left(E(x_i|y_i)\right)^2, \; i = m+1, \ldots, n. \tag{2}$$

In the **M-step**, we obtain the parameter estimator $\hat{\boldsymbol{\theta}}$ by maximizing the $E[l_c(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y})]$, which gives the explicit expression as

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\hat{\alpha},$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{m}(x_i - \hat{\alpha})\left[E\left(y_i|x_i; \boldsymbol{\theta}\right) - \bar{y}\right] + \sum_{i=m+1}^{n}\left[E\left(x_i|y_i; \boldsymbol{\theta}\right) - \hat{\alpha}\right](y_i - \bar{y})}{\sum_{i=1}^{m}(x_i - \hat{\alpha})^2 + \sum_{i=m+1}^{n}\left[E\left(x_i|y_i; \boldsymbol{\theta}\right) - \hat{\alpha}\right]^2},$$

$$\hat{\alpha} = \frac{1}{n}\left\{\sum_{i=1}^{m}x_i + \sum_{i=m+1}^{n}E\left(x_i|y_i; \boldsymbol{\theta}\right)\right\},$$

$$\hat{\delta}^2 = \frac{1}{n}\left\{\sum_{i=1}^{m}(x_i - \hat{\alpha})^2 + \sum_{i=m+1}^{n}\left[E\left(x_i^2|y_i; \boldsymbol{\theta}\right) - 2\hat{\alpha}E\left(x_i|y_i; \boldsymbol{\theta}\right) + \hat{\alpha}^2\right]\right\},$$

$$\hat{\sigma}^2 = \frac{1}{n}\left\{\sum_{i=1}^{m}\left[E\left(y_i^2|x_i; \boldsymbol{\theta}\right) - 2\left(\hat{\beta}_0 + \hat{\beta}_1 x_i\right)E\left(y_i|x_i; \boldsymbol{\theta}\right) + \left(\hat{\beta}_0 + \hat{\beta}_1 x_i\right)^2\right] + \sum_{i=m+1}^{n}\left[\left(y_i - \hat{\beta}_0\right)^2 + \hat{\beta}_1^2 E\left(x_i^2|y_i; \boldsymbol{\theta}\right) - 2\hat{\beta}_1\left(y_i - \hat{\beta}_0\right)E\left(x_i|y_i; \boldsymbol{\theta}\right)\right]\right\},$$

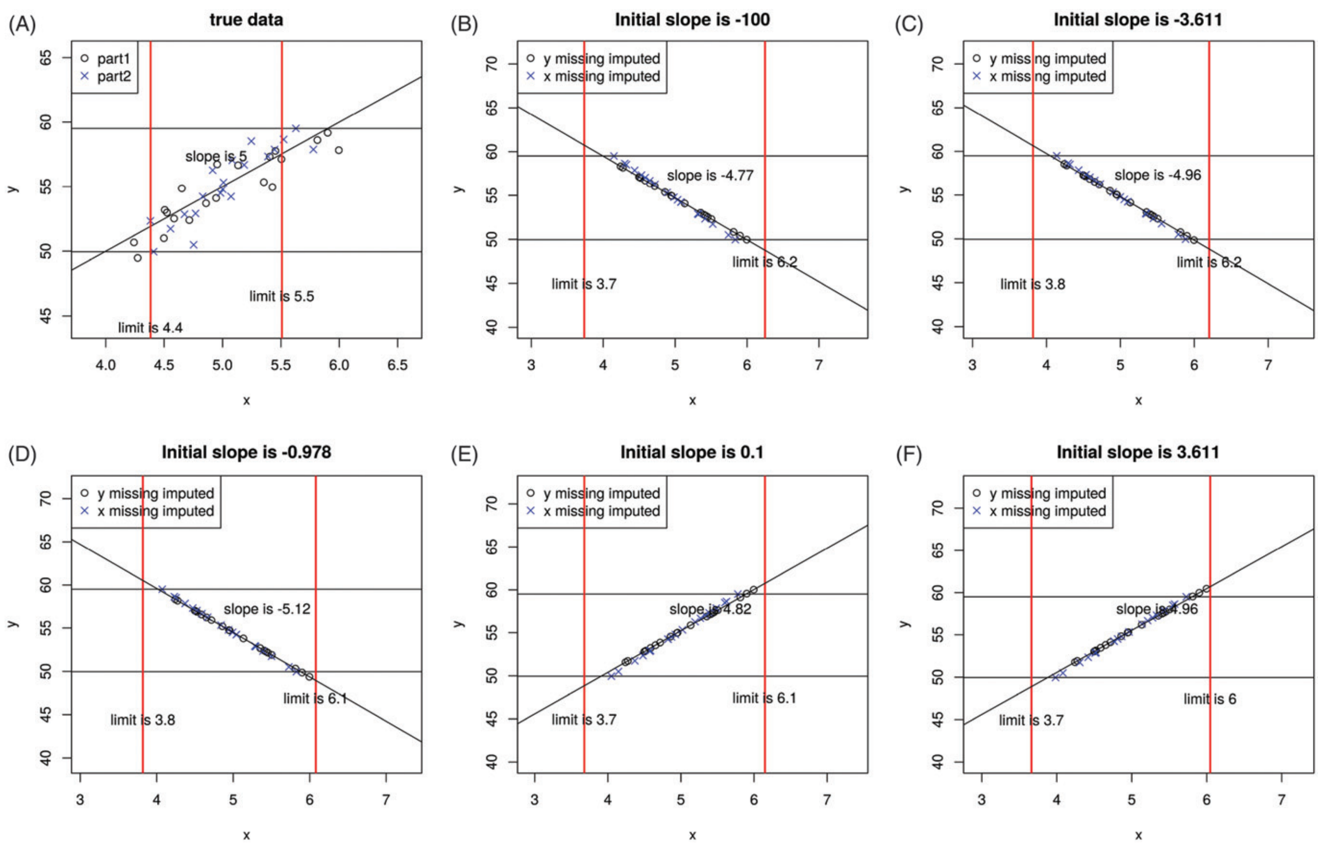

Figure 1. Simulation results of the estimated slope and the tolerance on the intermediate measurements under different scenarios: For A the complete data is used. From B to F the initial slope value increases from −100 to 3.611. The red vertical lines are the estimated tolerance on the intermediate measurements *x*.

where $\bar{y} = \frac{1}{n}\{\sum_{i=m+1}^{n} y_i + \sum_{i=1}^{m} E(y_i|x_i;\boldsymbol{\theta})\}$. The detailed derivations are in the Appendix.

It is noted that the initial β value could have a significant impact on the algorithm performance. When the absolute value of initial β_1 is large enough, the imputed complete data $(\boldsymbol{x}, \boldsymbol{y})$ can be fitted perfectly by the estimated model, making the range of the imputed $\boldsymbol{x}_{mis}$ values not overlapped with the range of $\boldsymbol{x}_{obs}$. That is, we can always have a perfect fitting on the unpaired data with a large value of β_1. When the absolute value of initial β_1 is small enough, the imputed $\boldsymbol{x}_{mis}$ values can be in a narrow region within the range of $\boldsymbol{x}_{obs}$, making the estimated parameters unstable. To address these issues, we add additional convergence conditions in the EM algorithm to ensure that the range of imputed $\boldsymbol{x}_{mis}$ values is partially overlapped with the range of observed $\boldsymbol{x}_{obs}$ values. Specifically, we start with an arbitrary initial β_1 to proceed with the E-step and the M-step, then check the condition whether the imputed $\boldsymbol{x}_{mis}$ values is partially overlapped with the range of observed $\boldsymbol{x}_{obs}$ values. We summarize the developed EM algorithm in Algorithm 1, which is implemented in the R package TOM and is available in Bitbucket (https://bitbucket.org/vtshen/tom/src/master/).

Algorithm 1

Initialize β_1, $\beta_0 = \bar{\boldsymbol{y}}_{obs}$, $\alpha = \bar{\boldsymbol{x}}_{obs}$, $\delta^2 = \frac{(\boldsymbol{x}_{obs}-\bar{\boldsymbol{x}}_{obs})^2}{m-1}$, $\sigma^2 = \frac{(\boldsymbol{y}_{obs}-\bar{\boldsymbol{y}}_{obs})^2}{n-m-1}$ and flag $=0$
repeat
  E-step: Compute $E[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})]$.
  M-step: Estimate β_0, β_1, α, σ^2 and δ^2 by maximizing $E[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})]$.
  if None of imputed missing values for $\boldsymbol{x}_{mis}$ are in the range of $\boldsymbol{x}_{obs}$, **then**
    $\beta_1 = \beta_1/2$, and flag $=1$.
  else
    if flag $=0$ **then**
    $\beta_1 = 2\beta_1$.
  end if
  end if
until all of β_0, β_1, α, σ^2 and δ^2 are converged.

The convergence of the EM algorithm can also be checked by the relative difference between the log-likelihood function. In the Appendix, we prove that the sign of the initial β_1 value does not affect the estimation of the tolerance of X under certain conditions.

Real data analysis

We analyze the real data in the open challenges (Jensen 2018), where the unpaired data $\boldsymbol{x}_{obs}$= (32.49, 34.07, 35.17, 35.04, 37.58, 37.72, 38.77, 40.07, 40.43, 41.92, 42.36, 42.4, 44.19, 45.07, 45.73, 46.54, 46.71, 48.21, 49.82, 51.27) and $\boldsymbol{y}_{obs}$= (26.56, 27.43, 28.17, 28.65, 28.77, 29.62, 29.9, 30.07, 30.13, 30.36, 30.51, 30.55, 30.97, 31.08, 31.53, 31.65, 31.9, 32.13, 32.31, 32.61). Table 1 shows the estimated parameters and tolerance of X under different initial slope values. The estimated slope is around 0.29 in the absolute value. For example, if one considers a positive slope, the estimated relationship between the intermediate measurements and final measurements could be $\hat{Y} = 18.44 + 0.29X$.

We suggest to choose a conservative estimation of the tolerance, $[a, b]$ by selecting the largest a and smallest b from Table 1. That is, the tolerance of the intermediate measurements can be [21.8, 59.0].

Simulation study

We conduct simulation by considering a linear regression situation,

$$Y = 3 + 5X + \epsilon, \ \ \epsilon \sim N\left(0, \sigma^2\right),$$

where $X \sim N(1, 0.5^2)$, X is independent of ϵ, and σ^2 is chosen such that the signal-to-noise ratio (Wu and Hamada 2009) is three. The total number of observations is 40. We remove the first 20 values in Y as missing, and remove the last 20 values in X as missing. Under the complete data, the tolerance of X is $[4.4, 5.5]$ given the specification of Y as $[50.0, 59.5]$ at a significant level $\tau = 1\%$.

Figure 1 shows the estimation performance of the proposed method under different initial slope values. Although different initial slopes lead to different model estimation, the estimation of the tolerance of X is quite similar and is close to the true tolerance.

Discussion

For problems with measurements being destructive, it is difficult to quantify the dependency between the intermediate measurements X and the final measurements Y. We proposed an EM-algorithm based method to accurately quantify the dependency, and obtain the tolerance of the intermediate measurements given the specification on the final measurements. The proposed method is not restricted to the simple linear regression between X and Y. It can be extended to a more general model as $Y = g(X) + \epsilon$ where $g(X)$ is a monotone function. The EM algorithm can be similarly developed.

Acknowledgments

The authors acknowledge the support from National Science Foundation CMMI-1435996 to this research.

ORCID

Xinwei Deng http://orcid.org/0000-0002-1560-2405

References

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1):1–38.

Horton, N. J., and N. M. Laird. 1999. Maximum likelihood analysis of generalized linear models with missing covariates. *Statistical Methods in Medical Research* 8(1):37–50.

Jensen, W. A. 2018. Open challenges: correlation of intermediate and final measurements. *Quality Engineering* 30(1):167–168.

Miljkovic, T., and N. Barabanov. 2015. Modeling veterans health benefit grants using the expectation maximization algorithm. *Journal of Applied Statistics* 42(6):1166–82.

Wu, C. F. J., and M. S. Hamada. 2009. *Experiments: Planning, Analysis, and Optimization*. 2nd ed., Hoboken, NJ: Wiley.

Appendix

Lemma 1. *Given the other estimated parameters β_0, α, σ_2, and δ_2, the sign of β_1 does not affect the tolerance of x.*

Proof. We have the conditional distribution of X given $Y = y$ as

$$X|Y = y \sim N\left(E\left(X|Y = y\right), Var\left(X|Y = y\right)\right),$$

where

$$E\left(X|Y = y\right) = \alpha + \frac{\beta_1\delta^2}{\sigma^2 + \beta_1^2\delta^2}\left(y - \beta_0 - \beta_1\alpha\right),$$
$$Var\left(X|Y = y\right) = \frac{\sigma^2\delta^2}{\sigma^2 + \beta_1^2\delta^2}.$$

In $E(X|Y = y)$, we substitute $E(y|x)$, equal to $(\beta_0 + \beta_1 x)$, for y, and obtain

$$\begin{aligned} E\left(X|Y = y\right) = & \quad \alpha + \frac{\beta_1\delta^2}{\sigma^2 + \beta_1^2\delta^2}\left(\beta_1 x - \beta_1\alpha\right) \\ = & \quad \alpha + \frac{\beta_1^2\delta^2}{\sigma^2 + \beta_1^2\delta^2}\left(x - \alpha\right). \end{aligned}$$

That is, $E(X|Y = y)$ does not depend on the sign of β_1. When the magnitudes of estimated parameters are the same, the tolerance of x keeps the same. ■

Derivation in the M-step

Note that the $E(l_c(\boldsymbol{\theta}; \boldsymbol{x}, \boldsymbol{y}))$ in the E-step is expressed as:

$$
\begin{aligned}
E_{\boldsymbol{x}_{mis},\boldsymbol{y}_{mis}}[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})] = \quad & -n\log(2\pi)-n\log(\delta\sigma)- \\
& \frac{1}{2\sigma^2}\sum_{i=1}^{m}\left[E\left(y_i^2|x_i;\boldsymbol{\theta}\right)-2(\beta_0+\beta_1x_i)E\left(y_i|x_i;\boldsymbol{\theta}\right)+(\beta_0+\beta_1x_i)^2\right]- \\
& \frac{1}{2\sigma^2}\sum_{i=m+1}^{n}\left[(y_i-\beta_0)^2+\beta_1^2E\left(x_i^2|y_i;\boldsymbol{\theta}\right)-2\beta_1(y_i-\beta_0)E\left(x_i|y_i;\boldsymbol{\theta}\right)\right]- \\
& \frac{1}{2\delta^2}\sum_{i=1}^{m}(x_i-\alpha)^2-\frac{1}{2\delta^2}\sum_{i=m+1}^{n}\left[E\left(x_i^2|y_i;\boldsymbol{\theta}\right)-2\alpha E\left(x_i|y_i;\boldsymbol{\theta}\right)+\alpha^2\right]
\end{aligned}
$$

Then in the M-step, we conduct parameter estimation of $\boldsymbol{\theta}$ by calculating the first order derivatives as follows.

1. Calculate the partial derivative of the $E(l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y}))$ with respect to the variable β_0:

$$
\begin{aligned}
0 \quad & =\frac{\partial}{\partial\beta_0}E_{\boldsymbol{x}_{mis},\boldsymbol{y}_{mis}}[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})] \\
& =-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\left[-2E\left(y_i|x_i;\boldsymbol{\theta}\right)+2(\beta_0+\beta_1x_i)\right]-\frac{1}{2\sigma^2}\sum_{i=m+1}^{n}\left[-2(y_i-\beta_0)+2\beta_1E\left(x_i|y_i;\boldsymbol{\theta}\right)\right] \\
\Rightarrow \quad & \sum_{i=1}^{m}\left[-2E\left(y_i|x_i;\boldsymbol{\theta}\right)+2(\beta_0+\beta_1x_i)\right]=\sum_{i=m+1}^{n}\left[2y_i-2\beta_0-2\beta_1E\left(x_i|y_i;\boldsymbol{\theta}\right)\right] \\
\Rightarrow \quad & \sum_{i=m+1}^{n}y_i+\sum_{i=1}^{m}E\left(y_i|x_i;\boldsymbol{\theta}\right)=n\beta_0+\beta_1\left[\sum_{i=1}^{m}x_i+\sum_{i=m+1}^{n}E\left(x_i|y_i;\boldsymbol{\theta}\right)\right].
\end{aligned}
$$

2. Calculate the partial derivative of the $E(l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y}))$ with respect to the variable β_1:

$$
\begin{aligned}
0 \quad & =\frac{\partial}{\partial\beta_1}E_{\boldsymbol{x}_{mis},\boldsymbol{y}_{mis}}[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})] \\
= \quad & -\frac{1}{2\sigma^2}\sum_{i=1}^{m}\left[-2x_iE\left(y_i|x_i;\boldsymbol{\theta}\right)+x_i(\beta_0+\beta_1x_i)\right]-\frac{1}{2\sigma^2}\sum_{i=m+1}^{n}\left[2\beta_1E\left(x_i^2|y_i;\boldsymbol{\theta}\right)-2(y_i-\beta_0)E\left(x_i|y_i;\boldsymbol{\theta}\right)\right] \\
\Rightarrow \quad & \sum_{i=1}^{m}\left[x_iE\left(y_i|x_i;\boldsymbol{\theta}\right)\right]+\sum_{i=m+1}^{n}y_iE\left(x_i|y_i;\boldsymbol{\theta}\right)= \\
& \left[\sum_{i=m+1}^{n}E\left(x_i|y_i;\boldsymbol{\theta}\right)+\sum_{i=1}^{m}x_i\right]\beta_0+\left[\sum_{i=m+1}^{n}E\left(x_i^2|y_i;\boldsymbol{\theta}\right)+\sum_{i=1}^{m}x_i^2\right]\beta_1.
\end{aligned}
$$

3. Calculate the partial derivative of the $E(l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y}))$ with respect to the variable α:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \alpha} E_{\boldsymbol{x}_{mis},\boldsymbol{y}_{mis}}[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})] \\
&= -\frac{1}{2\delta^2}\sum_{i=1}^{m}[2(x_i-\alpha)\times(-1)] - \frac{1}{2\delta^2}\sum_{i=m+1}^{n}[-2E(x_i|y_i;\boldsymbol{\theta})+2\alpha] \\
\Rightarrow n\alpha &= \sum_{i=1}^{m} x_i + \sum_{i=m+1}^{n} E(x_i|y_i;\boldsymbol{\theta}).
\end{aligned}
$$

4. Calculate the partial derivative of the $E(l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y}))$ with respect to the variable δ^2:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \delta^2} E_{\boldsymbol{x}_{mis},\boldsymbol{y}_{mis}}[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})] \\
&= -\frac{n}{2}\frac{1}{\delta^2} - \frac{1}{2}\sum_{i=1}^{m}\left[(x_i-\alpha)^2\right]\frac{-1}{\delta^4} - \frac{1}{2}\sum_{i=m+1}^{n}\left[E(x_i^2|y_i;\boldsymbol{\theta}) - 2\alpha E(x_i|y_i;\boldsymbol{\theta}) + \alpha^2\right]\frac{-1}{\delta^4} \\
\Rightarrow n\delta^2 &= \sum_{i=1}^{m}(x_i-\alpha)^2 + \sum_{i=m+1}^{n}\left[E(x_i^2|y_i;\boldsymbol{\theta}) - 2\alpha E(x_i|y_i;\boldsymbol{\theta}) + \alpha^2\right].
\end{aligned}
$$

5. Calculate the partial derivative of the $E(l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y}))$ with respect to the variable σ^2:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \sigma^2} E_{\boldsymbol{x}_{mis},\boldsymbol{y}_{mis}}[l_c(\boldsymbol{\theta};\boldsymbol{x},\boldsymbol{y})] \\
&= -\frac{n}{2}\frac{1}{\sigma^2} - \frac{1}{2}\sum_{i=1}^{m}\left[E(y_i^2|x_i;\boldsymbol{\theta}) - 2(\beta_0+\beta_1 x_i)E(y_i|x_i;\boldsymbol{\theta}) + (\beta_0+\beta_1 x_i)^2\right]\frac{-1}{\sigma^4} - \\
&\quad \frac{1}{2}\sum_{i=m+1}^{n}\left[(y_i-\beta_0)^2 + \beta_1^2 E(x_i^2|y_i;\boldsymbol{\theta}) - 2\beta_1(y_i-\beta_0)E(x_i|y_i;\boldsymbol{\theta})\right]\frac{-1}{\sigma^4} \\
\Rightarrow n\sigma^2 &= \sum_{i=1}^{m}\left[E(y_i^2|x_i;\boldsymbol{\theta}) - 2(\beta_0+\beta_1 x_i)E(y_i|x_i;\boldsymbol{\theta}) + (\beta_0+\beta_1 x_i)^2\right] + \\
&\quad \sum_{i=m+1}^{n}\left[(y_i-\beta_0)^2 + \beta_1^2 E(x_i^2|y_i;\boldsymbol{\theta}) - 2\beta_1(y_i-\beta_0)E(x_i|y_i;\boldsymbol{\theta})\right].
\end{aligned}
$$