Suppose we estimate the linear regression model *Z* = *α*0 *+ α*1*X*1 *+ α*2*X*2 *+ … + αnXn* where *Z* is fetal weight and *Y* is fetal age whereas *Xi* (s) are gestational ultrasound measures such as *bpd*, *hc*, *ac*, and *fl*. Suppose the random variable *Z* conforms normal distribution, according to equation 1 (Lindsten, Schön, Svensson, & Wahlström, 2017, pp. 8-9). Note, *Z* is random variable whereas *X* is data.

Where *α* = (*α*0, *α*1,…, *αn*)*T* is parameter vector and *X* = (1, *X*1, *X*2,…, *Xn*)*T* is data vector. The mean and variance of *Z* with regard to *P*(*Z* | *X*, *α*) are *αTX* and *σ*2, respectively. Suppose each has an inverse linear regression model *Xj* = *βj*0 *+ βj*1*Z.* In other words, *Zj* now is considered as the random variable conforming a normal distribution according to equation 2.

Where *βj* = (*βj*0, *βj*1)*T* is a partial parameter vector and (1, *Z*)*T* is a partial data vector. The mean and variance of each *Xj* with regard to the inverse distribution *Pj*(*Xj* | *Z*, *βj*) are *βjT*(1, *Z*)*T* and *σj*2, respectively. Of course, there are *n* inverse linear regression models.

Every missing value *zi* is estimated as the expectation based on the current parameter Θ*t*, according to equation 5.

Let *Ui* be a set of indices of missing values *xij*. In other words, if then, *xij* is missing. The set *Ui* can be empty. The equation 5 is written:

Note, *xi*0 = 1. According to equation 6, given *xij* is missing, we have:

Combining equation 5 and equation 6, we have:

It implies:

Equation 7 is used to estimate missing values *zi* in E-step.

|  |
| --- |
| 1. E-step: Missing values *zi* (s) are estimated according to equation 7, based on the current parameter Θ*t*.   Missing values *xij* (s) are estimated according to equation 6, based on the current parameter Θ*t*. |