Assignable Algorithms Available for Missing Data for Finding MV

P.Logeshwari
Research scholar, Department of Computer Science, NGM College Pollachi
Email: tppselvalogu@gmail.com
Dr.Antony Selvadoss Thanamani
,Associate Professor and Head, Department of Computer Science, NGM College Pollachi
Email: selvdoss@gmail.com

ABSTRACT

Assignable algorithms for use with missing data are becoming common- place in microcomputer packages. Specifically, 3 Assignable algorithms are currently available in existing software packages: the multiple-group approach, full information Assignable estimation, and the EM algorithm. Although they belong to this family of estimator, confusion appears to exist over the differences among the 3 algorithms. This article provides a comprehensive, nontechnical overview of the 3 Assignable algorithms. Multiple imputations, which is frequently used in conjunction with the EM algorithm, is also discussed.

Key word: Assignable algorithms, EM algorithm, multiple-group analysis, multiple imputation, software packages

I.INRODUCTION

Until recently ,the analysis of data with missing observations has been dominated by list wise (LD) and pair wise (PD) deletion methods (Kim & Curry, 1977; Roth,1994). However, alternative methods for treating missing data have become increasingly common in software packages, leaving applied researchers with a wide range of data analytic options. In particular, three maximum likelihood(ML) estimation algorithms for use with missing data are currently available: the multiple group approach(Allison,1987;Muthén,Kaplan,&Hollis,1987)canbeimplemented using existing structural equation modeling (SEM) software;Amos(Arbuckle,1995)andMx(Neale,1995)offerfullinformationmaximumlikelihood(FIML) estimation; and at least three packages, SPSS MissingValues,EMCOV(Graham&Hofer,1993),and NORM(Schafer,1998), incorporate the expectation maximization (EM) algorithm. The latter two programs alsooffermultipleimputation,asoutlinedbyRubin(1987).ThetheoreticalbenefitsofMLestimationarewidelyknown(Little&Rubin,1987),andsimulationstudieshavesuggestedthatMLalgorithmsmaybesuperiortotraditionaladhocmissingdatatechniquesinmanycases(Arbuckle,1996;Enders&Bandalos,inpress;Muthénetal.,1987;Wothke,2000).AlthoughmuchoftherecentmissingdataresearchhasbeenintheareaofSEM,agreatdealofconfusionapparentlyexistsoverthedifferencesamongthethreeMLmissingdataalgorithms.Forexample,asearchoftheSEMNETdiscussiongrouparchivesrevealedalargenumberofthreadsandrequestsforclarificationduringrecentyears,andthe frequency of these threadsdoesnotappeartobediminishing.Thatconfusionexistsisprobablynotasurpriseandiscertainlynotunwarranted;theMLalgorithmsappearfundamentallydifferentinmanyrespects,despitebelongingtothesameestimationfamily.AlthoughanextensivebodyoftechnicalliteratureexistsonMLmissingdatamethods(Dempster,Laird,&Rubin,1977;Finkbeiner,1979;Hartley&Hocking,1971;Little&Rubin,1987),nosinglereferenceisavailabletoappliedresearchersthatsuccinctlysummarizesthesimilaritiesanddifferencesamongthealgorithms.Thus,thegoalofthisarticleistoprovideathorough,nontechnicalprimeronthree widely available ML estimation algorithms for use with missing data: multiple group analysis, FIML, and the EMalgorithm. Multipleimputationalgorithms,whicharefrequentlyusedinconjunctionwiththeEMalgorithm,willalsobediscussed.

II.MULTIPLE-GROUP APPROACH

AnearlymethodforobtainingMLparameterestimatesinthepresenceofmissingdatawasgivenbyHartleyandHocking(1971).TheapplicationofthismethodtoSEManalyseswasoutlinedbyAllison(1987)andMuthénetal.(1987)andhassincebeenreferredtoasthemultiplegroupmethod.Inthisprocedure,asampleisdividedinto*G*subgroups,suchthateachsubgrouphasthesamepatternofmissingdata.Thatis,observationswithineachofthe*G*subgroupshavethesamesetofvariables present and missing. A likelihood function is computed for each of the *G* groups, and the group wise likelihood functions are accumulated across the entire sample and maximized. Although mathematically unrelated,thisalgorithmislooselyanalogoustoPD;asubgroup*gi* contributestotheestimationofallparametersthatinvolvetheobserveddatapointsforthatgroupbutdoesnotcontributeto parameters that involve missing-data points. Assumingmultivariatenormality,theloglikelihoodfunctiongivenbyHartley and Hocking (1971) is

$$-1/2 \sum_{g=1}^{G} ng[\log|\Sigma_g| + tr(sg\,\Sigma_g^{-1}) + tr(Hg\,\Sigma_g^{-1}) + Cg]$$

where $H_g=(x_g-\mu_g)(x_g-\mu_g)'$.For each of the G sub groups, $ng$is the number of observations,Σg and S$g$aretheparameterestimatesandsamplemoments,respectively ,Cg is a constant that depends on the data,andHgcontainsthevectorofmeanresiduals.Becausethe$G$subgroupshavedifferentpatternsofmissingdata,thisimpliesthattheelementsofx_g,μ_g,S_g,and$\Sigma_g$andaredifferentforeachgroup.To illustrate,considerasimplemodelcomprisingthreeobservedvariables:*X1*,*X2*,and *X3*. Furthermore, suppose a subgroup, *g1*, has complete data on *X1* and *X3*,butismissing*X2*.The$\mu_g$andΣgtermsinthegroupwiselikelihoodfunctionfor*g1*would contain only the parameter estimates that involve *X1* and *X3,* as follows:

$$\mu=[\mu_1\text{-}0\text{-}\mu_3] \text{ and } = \begin{bmatrix} \sigma_{11} & 0 & \sigma_{13} \\ 0 & 0 & 0 \\ \sigma_{31} & 0 & \sigma_{33} \end{bmatrix}$$

Similarly,$x_g$andSgwouldcontainthecorrespondingsamplemomentstakenfrom the ngcomplete observations in $g1$.Allison (1987) and Muthén et al. (1987) demonstrated how to implement HartleyandHocking's(1971)algorithmusingtheLISRELmultiple-groupspecification, which maximizes the likelihood equation.

$$-1/2 \sum_{g=1}^{G} ng[\log|\textstyle\sum_g| + tr(sg\sum_g^{-1}) + Cg]$$

This function is clearly similar to Equation 1, but does not include a term for the vector of mean residuals—LISREL does allow for the addition of ameanvectorterm,however.IntheusualSEMmultiplegroupanalysis,*G*groupsareformedthatrepresentindependentlysampledsubpopulations(e.g.,menandwomen),anditistypicallyofinteresttodeterminewhethersomespecifiedsetofparametersorparameter values are common to the *G* groups. In the missing-data application, the subpopulations correspond to the *G* patterns of missing data required by Hartley and Hocking's algorithm. The additionalinformationfromthegroupswithpartiallyrecordeddataisincorporatedbythespecificationofparameterequalityconstraints across the *G* groups.Despite the wide availability of the LISREL program at the time, the multiplegroupmethodofmissingdataanalysishadpracticallimitationsthatpreventeditswidespreaduse.AspointedoutbyArbuckle(1996),theLISRELspecificationforthemultiplegroupapproachrequiredanexceptionallevelofexpertiseandthus was practically limited to situations in which there are only a small number of missing-data patterns. Muthénetal.(1987)andKaplan(1995)describedsituationsinwhichthismightoccur(e.g.,BIBspiraleddesigns),butthenumberofdistinctmissing-data patterns is often quite largeinappliedsettings,makingthemethoddifficulttoimplement.Despitethetechnicaldifficultiesassociatedwithitsimplementation,themultiple group approach does have advantages. First, the method can be used to estimate both just-identified (e.g., correlation, regression) and overidentified(e.g.,SEM)modelparameters.ThisisapointofcontrastwiththeEMalgorithm,whichcannotcurrentlybeusedtodirectlyestimatelinearmodelparameters.Second,itis important to note that the multiple-group approach doesnotestimate,orimpute,missingobservations,butyieldsdirectestimatesofmodelparametersandstandarderrors.Thisisanadvantage,asadditionalcorrectiveproceduresarenotnecessary to obtain standard error estimates. Third, the multiple-group approach yields the usual chi-square test statistic for model fit, although the degrees of freedom and accompanying *p* value are incorrect due to the use of dummy values in the input covariance matricesofsubsampleswithmissingvariancecovarianceelements.However,thisiseasilyremediedbysubtractingthenumberofpseudovaluesfromthedegreesoffreedomterm.Finally,asabyproductofthemultiplegroupspecification,thechisquarestatisticcanalsobeusedtotesttheMCARassumption.IftheMCARassumptionholds,parameterestimatesacrosssubgroupsshouldbeequal.Thus,thechisquaredifferencetestoftheequalityconstraintsimposed acrossthe *G* subgroups is also a test of the MCAR assumption; a statistically significant χ^2 value suggests that data are not MCAR.

III.FIML

Twostructuralequationmodelingsoftwarepackagescurrentlyofferfimlestimationroutinesformissingdata:AMOS(Arbuckle, 1995)andMx(Neale,1995).The FIML approach was originally outlined by Finkbeiner (1979) forusewithfactoranalysisandissimilartothemultiplegroupmethod,exceptthatalikelihoodfunctioniscalculatedattheindividual, ratherthanthegroup,level.Forthisreason,the FIML approach has been referred to as raw maximum likelihood estimation (Duncan, Duncan, & Li, 1998; Graham, Hofer, & MacKinnon, 1996).

Like the multiple-group approach, the FIML algorithmisconceptuallyanalogoustoPD(althoughmathematicallyunrelated)inthesensethatallavailabledata is used for parameter estimation. An examination of the individuallevellikelihoodfunctionillustratesthispoint.Assumingmultivariatenormality,thecase wise likelihood of the observed data is obtained by maximizing the function

$$\log L_i = K_i - 1/2 \log|\textstyle\sum_i| - 1/2\log(x_i - \mu_i)' \sum_i^{-1}(x_i - \mu_i)$$

where$x_i$isthevectorofcompletedataforcase$i$,$\mu_i$containsthecorrespondingmeanestimatesderivedfromtheentiresample,and$K_i$is aconstantthatdependsonthenumberofcompletedatapointsforcase$i$.Likeμ_i,thedeterminantandinverseof$\Sigma_i$arebasedonlyontho sevariablesthatareobservedforcase*i*.Theoveralldiscrepancy function value is obtained by summing the *n* case wise likelihood functions as follows:

$$\log L(\mu\text{-}\sum)=\sum_{i=1}^{N}\log L_i$$

Toillustrate,supposeMLparameterestimatesaresoughtforamodelcomprisedof three observed variables: *X1*, *X2*, and *X3*. The parameters of interest are

$$\mu=[\mu_1\text{-}\mu_2\text{-}\mu_3] \text{ and } =\begin{bmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13}\\ \sigma_{21} & \sigma_{22} & \sigma_{23}\\ \sigma_{31} & \sigma_{32} & \sigma_{33}\end{bmatrix}$$

Thelikelihoodvalueforanobservationwith*X2*missingwouldbeafunctionofthetwocompleteobservationsaswellastheparameterestimatesthatinvolved*X1*and *X3*. The relevant parameters are shown in the following.

$$\mu=[\mu_1\text{-}0\text{-}\mu_3] \text{ and } \sum =\begin{bmatrix}\sigma_{11} & 0 & \sigma_{13}\\ 0 & 0 & 0\\ \sigma_{31} & 0 & \sigma_{33}\end{bmatrix}$$

Basedonthepreviousexamples,themathematicalsimilaritiesbetweenthemultiplegroupandFIMLalgorithmsshouldbeapparent; theprimarydifferenceisthat FIMLfittingfunctionisthesumof*n*casewiselikelihoodvalues,whereasthemulti- plegroup function is the sum of *G* group wise likelihood values.SeveralpointsshouldbemadeabouttheFIMLalgorithm. First,likethemultiplegroupapproach,oneoftheadvantagesoftheFIMLalgorithmisitsapplicability to both just-identified and over-identified models. In the latter case, thelikelihoodequationinEquation3isextendedsuchthatthefirst andsecondordermoments(μandΣ,respectively)areexpressedas functionsofsomeparametervector,*γ*(Arbuckle,1996).Assuch,themethodisquitegeneralandcanbeappliedtoawidevarietyofanalyses,includingtheestimationofmeans,covariancematrices,multipleregression,andSEM.Second,whenusedinSEMapplications,FIMLyieldsachisquaretestofmodelfit.However,thechisquarestatisticgeneratedby FIML does not take the usual form *F*(*N* – 1), where *F* isthevalueofthefittingfunction.Clearly,thechisquaretestcannotbecalculatedinthenormalfashion,asthereisnosinglevalueof*N*thatisapplicabletotheentiresample.Also,unliketheusualSEMfittingfunctions,thereisnominimumvalueassociatedwiththeFIML log-likelihood function, although the value of this statistic will increaseasmodelfitworsens.Instead,achisquaretestformodelfitiscalculatedasthedifferenceinloglikelihoodfunctionsbetweentheunrestricted(*H*0)andrestricted(*H*1)moelswithdegreesoffreedomequaltothedifferenceinthenumberofestimatedparameters between the two models. Third, although many popular fit indexes canbecomputedunderFIML,thespecificationofameansstructure(requiredforestimation)renderscertainfitindexesundefined(e.g.,GFI).Fourth,similartoPD,indefinitecovariance matrices are a potential byproductof the FIML approach. However, Wothke (2000) suggested thatindefinitenessproblemsarelesspervasivewithFIMLthanwithPD.Fifth,unliketheEMalgorithm(discussedinthefollowing), standard error estimates are obtained directly fromtheanalysis,andbootstrappingisnotnecessary.Finally,itisimportanttonotethattheFIMLalgorithm does not impute missing values; only model parameters are estimated.

IV.EM ALGORITHM

At least three packages currently implement the EMalgorithm:SPSSMissingValues,EMCOV(Graham&Hofer,1993),andNORM(Schafer,1998).An early work by Orchard and Woodbury (1972) explicated the underlying method, which they called the “missing informationprinciple.”Dempsteretal.(1977)providedanextensivegeneralizationand illustrationof the methodand namedittheEMalgorithm.TheEMalgorithmusesatwostepiterativeprocedurewheremissingobservationsarefilledin,orimputed,andunknownparametersaresubsequentlyestimated.Inthefirststep(the*E*step),missingvaluesarereplacedwiththe conditional expectation of the missing data giventheobserveddataandaninitialestimateofthecovariancematrix.Thatis,missingvaluesarereplacedbythepredictedscoresfromaseriesofregressionequationswhereeachmissingvariableisregressedontheremainingobservedvariablesforacase*i*.Usingtheobservedand imputed values, the sumsandsumsofsquaresandcrossproductsarecalculated.Toillustrate,supposeameanvectorandcovariancematrix,θ=(μ,Σ),issoughtforan*n*×*K*datamatrix,*Y*,thatcontainssetsofobservedandmissingvalues(Y_{obs}andY_{mis},respectively).Usingtheobservedvalues(Y_{obs})andcurrentparameterestimates($\theta^{(t)}$),thecalculations forthesufficientstatisticsatthe*t*thiterationofthe*E* step are

$$\sum_{i=1}^{n} y_{ij}\,|\, y_{obs},\theta^{(+)} =\sum_{i=0}^{n} y_{ij}^{(t)} \quad j=1,\ldots\ldots\ldots k$$

$$\sum_{i=1}^{n} y_{ij}\, y_{ik} \,|\, y_{obs},\theta^{(+)}=\sum_{i=0}^{n} y_{ij}^{(t)} y_{ik}^{(t)} c_{jki}^{(t)} j,k=1,\ldots k$$

where

$$y_{ij}^{(t)} = \begin{cases} y_{ij} , \sum(y_{ij}\,|\, y_{obs} ,\theta^{(t)}), & \text{if } y_{ij} \text{ is observed} \\ & \text{if } y_{ij} \text{ is missing}\end{cases}$$

and

$c_{ikj}^{(t)} = \text{cov}\left[\left(y_{ij}, y_{ik} \mid y_{obs}, \boldsymbol{\theta}^{(t)} \right) \right.$,if y_{ij} or y_{ik} is observed,
if y_{ij} and y_{ik} are missing

Thus, missing values of *yij*are replaced with conditional means and covariance's given the observed data and the current set of parameter estimates.2 It should be noted that the preceding formulas can be found in Little and Rubin (1987).Inthesecondstep(the*M*step),MLestimatesofthemeanvectorandcovariancematrixareobtainedjustasiftherewerenomissingdatausingthesufficientstatistics calculated at the previous *E* step. Thus, the *M* step is simply a complete-data ML estimation problem. The resulting covariance matrix and regressioncoefficientsfromthe*M*steparethenusedtoderivenew estimatesofthemissingvalues·AspointedoutbyLittleandRubin (1987),missingvaluesarenotnecessarilyreplacedwithactual data points, but are replaced by the condition functions of the missing values in the complete-data log-likelihood.at the next *E* step, and the processbeginsagain.Thealgorithmrepeatedlycyclesthroughthesetwostepsuntilthedifferencebetweencovariancematricesinsubsequent *M* steps falls below some specified convergence criterion. Readers are encouraged to consult Little and Rubin (1987) for further technical details.Several points should be noted concerning the EM algorithm. First, unlike the multiple-group and FIML approaches, the EM algorithm cannot be used to obtain direct estimates of linear model parameters (e.g., regression, SEM); as currently implemented, the EM algorithm can only be used to obtain ML estimates of a mean vector and covariance matrix. Obviously, this matrix can be used for input in subsequent linear model analyses. Additionally, the covariance matrix can be used to estimate,or impute,missing-datapoints at thefinaliteration.The latter approach may, at first glance, be appealing due to the illusion of a complete data set, but there is a notable drawback associated with thispractice. Al- though the imputed values are optimal statistical estimates of the missing observations, they lack the residual variability present in the hypothetically complete data set; the imputed values fall directly on a regression line and are thus imputed without a random error component. As a result, standard errors from subsequent analyses will be negatively biased to some extent, and bootstrap (Efron, 1981) procedures must be employedtoobtaincorrectestimates.Alternatively,multipleimputationproceduresdesignedtorecoverresidualvariabilityareavailableintheEMCOV(Graham&Hofer,1993)andNORM(Schafer,1998) packages and are discussed next. However, it is important to note that a correction factor is added to the conditional expectation of the missingdataateach*E*steptocorrectforthisnegativebiasintheoutputcovariancematrix;thisisseeninthec_{jkl}Equation5.Although nostudieshavecomparedtheimpactofthesetwoEMmethodsinthecontextofSEM,itseemsreasonabletorunanalysesusingtheoutputcovariancematrixratherthanthesinglyimputeddataset.Despitethedifficultiespreviouslynoted,theEMalgorithmmaybepreferredinsituationswherethemissingdatamechanism(i.e.,thevariablesareassumedtoinfluence messiness) is not included in the linear model being tested. This is be- cause the MAR assumption discussed previously is defined relativetotheanalyzedvariablesinagivendataset.Forexample,ifthemissingvaluesonavariable*Y*aredependentonthevaluesofanothervariable*X*,theMARassumptionno longer holds if *X* is not included in the ultimateanalysis.Thisisclearlyproblematicforthetwodirectestimationalgorithms,as*X*mustbeincorporatedinthesubstantive model for MAR to be tenable. However, this is not the case with the EM algorithm, as the input covariance matrix used to estimatesubstantivemodelparametersmaybeasubsetofalarger covariancematrixproducedfromanEManalysis. In this case, the EM mean vector and covariancematrixareestimatedusingthefullsetofobservedvariables,andtheelementsthatareofsubstantiveinterestareextractedforfutureanalyses.Ofcourse,theapplicationoftheEMalgorithminthisscenarioassumesthattheresearcherhasexplicitknowledgeofthemissing-data mechanism, which may not likely be the case in practice. Nevertheless, the use of the EM algorithm in the manner described previously may make the MAR assumption more plausible in certain circumstances.

V.MULTIPLE IMPUTATION

TheprimaryproblemassociatedwithEMalgorithmisthatthevariabilityinthehypotheticallycompletedatasetisnotfullycaptured duringtheimputationprocess.Multipleimputation,asoutlinedbyRubin(1987),creates*m*>1imputeddatasetsthatareanalyzedusingstandardcompletedatamethods.The*m*setsofparameterestimatesaresubsequentlypooledintoasinglesetofestimatesusingformulaspro- vided by Rubin. The logic of multiple imputation is based on the notion thattwosourcesofvariabilityarelostduringtheEMimputationprocess.Asdescribedpreviously,thefirstoccursduetoregressionimputation;imputedvaluesfalldirectlyontheregressionlineandthuslackresidualvariability.Thesecondsourceoflost variability is due to the fact that the regression equationsarederivedfromacovariancematrixthatis,itself,estimatedwitherrorduetothemissingdata.Thatis,thecovariancematrixusedtoimputevaluesisoneofmanyplausiblecovariance matrices. The multiple imputation process attemptstorestorethelostvariabilityfrombothofthesesources.Currently,thereareatleasttwowidelyavailablemultipleimputationprogramsbasedontheEMalgorithm:EMCOV(Graham&Hofer,1993) and NORM (Schafer, 1998).[3]Althoughconceptuallysimilar,themultipleimputationalgorithmsarequitedifferent:EMCOVgenerates*m*imputeddatasetsusingthebootstraptechnique,whereasNORMdoessousingBayesiansimulation.Following an initial EM analysis, EMCOV (Graham & Hofer, 1993) restores residual variability by adding a randomly sampled (with replacement) residual term to each of the imputed data points. For every nonmissing-data pointin theoriginaldataset,avectorofresidualsforeachvariableiscalculatedasthedifference between the actual and predicted values from the regression equations (all othervariablesservingaspredictors)usedtoimputemissingvalues.Next,*m*datasetsarecreatedbyrepeatedlyimputingmissingvaluestotheoriginaldatasetsuchtht*m*1imputationsarebasedonnewestimatesofthecovariancematrix.Inthefirststep,abootstrapisperformedontheoriginaldata,yieldinganewdatamatrix of the same dimensions as the original. Next,thebootstrappeddataareanalyzedusingtheEMalgorithm, andanewestimateofthecovariancematrixisobtained.Finally,missingvaluesintheoriginaldatasetareimputedusingregressionequationsgeneratedfromthenewcovariancematrix.Thisbootstrap processisrepeated*m*1times(theimputeddatamatrixfromtheoriginalEManalysisservesasthefirst of the *m* data sets), and residual variation is restored to the*m*1setsofimputeddatapointsusingrandomlysampledresidualterms,asdescribedpreviously.Incontrast,NORM(Schafer,1998)usesiterativeBayesiansimulationtogenerate*m*imputeddatasets.LiketheEMalgorithm,theNORMalgorithmrepeatedly cycles through two steps: Missing observationsareimputed(theimputation,or*I*step)andunknown parametersareestimated(theposterior,or*P*step).However,unlikeEM,thedataaugmentation(DA)algorithmimplemented in NORM uses a stochastic rather than a deterministicprocess.Inthefirststep,missingdatapointsarereplacedbyrandomlydrawnvaluesfromtheconditionaldistributionofthemissingdatagiventheobserveddataandacurrentestimateoftheparametervectorθ;parameterestimatesfromanEManalysisprovidestartvaluesforthefirstiteration.Next,newparameterestimatesarerandomlydrawnfromaBayesianposteriordistributionconditionedontheobservedandimputedvaluesfromthefirststep.Thesenewparametervaluesareusedtoimputevaluesinthesubsequent*I*step,andtheprocessbeginsagain.Thistwosteppr ocedureisiterateduntilconvergenceoccurs,atwhichpointthefirstof*m*imputeddatamatricesiscreatedfromafinal*I*step.AdditionalimputeddatasetsareobtainedbyrepeatingtheDAprocess*m*1 times.Finally,itshouldbenotedthatthestochastic nature of the DA process requires a different convergence criterion than theEMalgorithm.BecauseDAparameterestimatesaredrawnrandomlyfromaposteriorprobabilitydistribution,valueswillnaturallyvarybetweensuccessiveiterations, even after convergence occurs. Thus, the DA algorithm converges when the *distribution* of the parameter estimates no longer changes between contiguousiterations.ReadersareencouragedtoconsultSchafer(1997)andSchaferand Olsen (1998) for further details.After implementing EMCOV or NORM, complete-data analyses are per- formed on each of the *m* imputed data sets, and the parameter estimates from these analyses are stored in a new file. Using rules provided by Rubin (1987), a single set of point estimates and standard error values can be obtained; both EMCOV and NORM include routines that will perform the necessary calculations. Two final points should be made regarding multiple imputation. First, Schafer (1997) suggested that adequate results could be obtained using as few as five imputed data sets. Second, a straightforward method of obtaining SEM goodness-of-fit tests is not currently available, although work on the topic is on- going(Schafer&Olsen,1998).

VI.SUMMARY

Recent software advances have provided applied researchers with powerful options for analyzing data with missing observations. Specifically, three MLalgorithms(multiplegroupanalysis,FIML,andtheEMalgorithm)arewidelyavailable in existing software packages. However, the wide array of dataanalyticoptionshasresultedinsomeconfusionoverthedifferencesamongthethreealgorithms.Assuch,thegoalofthisarticlewastoprovideabriefoverviewofMLalgorithmsinhopesthatappliedresearcherscanmakeinformeddecisionsregardingthe use of ML algorithms in various data analytic settings. the EM algorithm may be preferable when the missing-data mechanism does not appear in the substantive model.

REFERENCES

[1]Allison,P.D.(1987).Estimationoflinearmodelswithincompletedata.InC.C.Clogg(Ed.),*Sociologi- cal methodology, 1987* (pp. 71–103). San Francisco: Jossey-Bass.

[2]Arbuckle, J. L. (1995). Amos user’s guide [Computer software]. Chicago: Smallwaters.

Arbuckle,J.L.(1996).Fullinformationestimationinthepresenceofincompletedata.InG.A.

[3]Marcoulides&R.E.Schumacker(Eds.),*Advancedstructural equationmodeling*(pp.243–277). [4]Mahwah, NJ: Lawrence Erlbaum Associates, Inc.Dempster,A.P.,Laird,N.M.,&Rubin,D.B(1977).Maximumlikelihoodfromincompletedataviathe EM algorithm.*Journal of the Royal Statistical Society, Ser. B, 39,* 1–38.

[5]Duncan,T.E.,Duncan,S.C.,&Li,F.(1998).Acomparisonofmodel-andmultipleimputation-based approachestolongitudinalanalyseswithpartialmissingness.*StructuralEquationModeling,5,*1–21.

[6]Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap, and other resampling methods. *Biometrika, 68,* 589–599.

[7]Enders,C.K.,&Bandalos,D.L.(inpress).Therelativeperformanceoffullinformationmaximumlike- lihood estimation for missing data in structural equation models.*Structural Equation Modeling*.

Finkbeiner,C.(1979).Estimationforthemultiplefactormodelwhendataaremissing.*Psychometrika*,

44, 409–420.
[8]Graham,J.W.,&Hofer,S.M.(1993).EMCOVreferencemanual[Computersoftware].LosAngeles: University of Southern California, Institute for Prevention Research.
[9]Graham,J.W.,Hofer,S.M.,&MacKinnon,D.P(1996).Maximizingtheusefulnessofdataobtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research, 31,* 197–218.
[10]Hartley, H. O., & Hocking, R. R. (1971).The analysis of incomplete data.*Biometrics*, *27,* 783–823.
[11]Kaplan,D.(1995).TheimpactofBIBspiraling-inducedmissingdatapatternsongoodness-of-fittestsin factor analysis. *Journal of Educational and Behavioral Statistics, 20*, 69–82.
[12]Kim,J.,&Curry,J.(1977).Thetreatmentofmissingdatainmultivariateanalyses.*Sociological Methods & Research*, *6,* 215–240.