Am J Hum Genet 39:775–786, 1986

A General Multivariate Approach to Linear Modeling in Human Genetics

GREGORY CAREY[1]

SUMMARY

The general linear structural equation model is applied to problems in human genetics where there may be more than one measured phenotype per individual. A modeling convention, termed conditional associations, is developed to extend the general linear model so that it can handle the unique problems in human genetic models that arise from the pairing up of individuals or families under assortment between mates and the assortative placement of adoptees. Formulas are presented to generate expected covariance matrices for assortment or assortative placement on many variables simultaneously. It is demonstrated that all linear models in human genetics can be reduced in form to two fundamental equations. An algorithm is presented that will allow the application of these two equations to linear modeling in human genetics.

INTRODUCTION

The past decade has seen an increase in the application of linear modeling or path analysis to problems in human genetics [1–7]. Multivariate models (i.e., models with more than one phenotypic measure per individual) have been developed for twin data [8], and bivariate path models have been presented for nuclear families [9]. Recently, Vogler [10] gave general matrix equations for several models in familial resemblance on certain types of kinships.

Received March 25, 1986; revised June 10, 1986.

Part of this research was done while G. C. was at the Department of Human Genetics, Medical College of Virginia, Richmond, Va., and was supported there in part from grants GM-30250, GM-32732, and AG-04954. This research was also supported at the University of Colorado by grant HD-07289 from NICHD and a grant from the Spencer Foundation to J. C. DeFries.

[1] Institute for Behavioral Genetics and Department of Psychology, Campus Box 447, University of Colorado, Boulder, CO 80309.

Linear models have also been extended to encompass phenomena such as assortative mating. Cloninger [11] developed the notion of the copath, and Van Eerdewegh [12] presented an analogous but general concept of a delta path. Van Eerdewegh's formulation has the advantage of starting with first principles and explicit assumptions. He presents general matrix formulas that one can apply to multivariate assortative mating and assortative placement.* Cloninger's copath formula has not been formally generalized to the multivariate case, and his tracing rule may give erroneous expectations with some models [12].

Here, the general linear structural equation model [13] is applied to multivariate problems in genetics and a rationale is developed for extending the general linear model to account for multivariate assortative mating and assortative placement.

THE GENERAL LINEAR MODEL

Let Y denote a vector of endogenous variables. In a path diagram, vector Y consists of all variables with a straight arrow going into them. Let Z denote the vector of all other variables in the model. Z is often subdivided into exogenous and residual variables, usually denoted, respectively, as X and U. Let W be a matrix of weights or path coefficients with subscripts determining the appropriate vectors. For example, W_{YX} is the matrix of path coefficients from the Y variables to the exogenous variables. The characteristic structural equation is

$$Y = W_{YY}Y + W_{YZ}Z \ . \tag{1}$$

Let C symbolize a covariance matrix with subscripts denoting the variables involved in the matrix, and let R signify a correlation matrix. The covariance matrix for all variables in the model is

$$\begin{array}{c} \\ Z \\ Y \end{array} \begin{array}{c} \begin{array}{cc} Z & Y \end{array} \\ \begin{pmatrix} C_{ZZ} & C_{ZZ}B' \\ BC_{ZZ} & BC_{ZZ}B' \end{pmatrix} \end{array} , \tag{2}$$

where $B = (I - W_{YY})^{-1}W_{YZ}$. Predicted covariances are a function of the covariance matrix among the Z variables and the two weight matrices. Consequently, by knowing C_{ZZ}, W_{YY}, and W_{YZ} predicted covariances for all variables in the model are known.

* An anonymous referee pointed out that the term "selective placement" has been historically used to refer to assortative placement of adoptees with adoptive families, not to selection of adoptees and their families in the strict sense of the term "selection." I recognize this distinction and use the term "assortative placement" throughout this paper. The unqualified term "assortment" will refer to marital assortment.

An important corollary of expression (2) is: If $Y_1 = W_{11}A_1 + W_{12}A_2 + \ldots W_{1m}A_m$ and $Y_2 = W_{21}A_1 + W_{22}A_2 + \ldots W_{2n}A_n$, then

$$C_{Y1Y2} = \sum_{i=1}^{m} \sum_{j=1}^{n} W_{1i} C_{AiAj} W_{2j}' \quad . \tag{3}$$

Here A denotes vectors that are causally antecedent to Y. Equation (3) applies only to recursive linear models.

Subsets of Y variables may be recoded as Z variables by deriving C_{YZ} and C_{YY} and augmenting C_{ZZ} by these two blocks. This recoding must be done in proper order. Define the *level* of a Y variable as the maximum number of causal arrows found by tracing backwards from the variable to a Z variable. Then Y variables at levels 1 through n may be recoded as Z variables with respect to the remaining Y variables at levels (n + 1) through (n + m), $m > 0$.

The general linear model (GLM) in equations (1), (2), and (3) has been used to solve for predicted covariances in a number of disciplines such as psychometrics, sociology, and econometrics [13–16]. There are two problems, however, with its direct application to human genetics. First, assortative mating and assortative placement often express elements of the covariance matrix among the Z variables in terms of correlations or covariances among the endogenous variables. Second, constraints, such as the assumption that gene-environment correlation is at equilibrium over generations, may also result in the expression of elements of one of the parameter matrices as a function of C_{YY} or C_{YZ}.

Still, the parameterization in equations (1), (2), and (3) suggests a method of solving for predicted covariance matrices in genetic problems in the presence of assortment and assortative placement. If one could derive the effects of assortment or assortative placement on the covariance matrix among the Z variables, then one could substitute this new C_{ZZ} matrix into the equations and calculate the expected covariances for the rest of the model.

CONDITIONAL ASSOCIATIONS: ASSORTATIVE MATING AND ASSORTATIVE PLACEMENT

Traditional path analysis uses two conventions to represent the relationship between pairs of variables. A single-headed arrow between two variables signifies causality. A double-headed arrow may join two Z variables, say A and B, and denotes an unspecified source of covariance; that is, A may cause B; B may cause A; other variables may jointly cause A and B; or some combination of the above may occur. Here, we introduce a different concept, related to the copath [11] and the delta path [12], that models the effects of a sort/merge process on the covariance structure of linear models.

First define a *set* as an individual or collection of individuals and *all* the variables of the individual(s) that temporally occur up to the time of assortment. Assortment and assortative placement are modeled as an imperfect sort/merge procedure between two sets that are otherwise independent. Examples of such paired sets are {husbands} and {wives} in assortment or {biological

mothers and biological fathers} and {adoptive mothers and adoptive fathers} in assortative placement. By these definitions, the pairing up is done between *sets*. Variables are included within each set. Some variables form the basis for pairing, but actual pairing is done between concrete individuals, families, etc. Of course, this process is a figurative analytical tool; it does not necessarily imply a physical, spatiotemporal sorting of, say, men and women.

Sorting does not change the value of any variables; all means, variances, and covariances remain unchanged *within* each set. Thus, GLM may be used to derive the covariance matrix within each set. Denote the variables as vectors A and A* for set one and set two, respectively. For example, if assortment is modeled, then genotypes (G), environments (E), and phenotypes (P) for husbands may constitute A; and G, E, and P for wives, the variables in A*. Equations (1) and (2) or (3) allow solution for covariances within each set, or C_{AA} and $C_{A^*A^*}$. Because Y variates can be recoded as Z variates, all elements in A and A* may be regarded as Z variables.

To model assortment, let M denote a subset of A that are the primary assortment variables. For example, if assortment is based on both phenotype and environment, then M = (E, P)′ for the father. Let M* be an analogous subset for A*. Assume that all variables are standardized; the unstandardized case is given later. Let D denote a matrix of assortment parameters between M and M* such that d_{ij} is that part of the correlation between variable i, $i \in$ M, and variable j, $j \in$ M*, that is due solely and exclusively to a matching on these two variables and these two variables alone. d_{ij} is a parameter and need not be the observed correlation between the two variables.

Parameter d_{ij} can be modeled after Carey and Rice's [1] derivation in the case of univariate phenotypic assortment—conditioning the distribution of one mate's variables on the other mate's phenotype. To generalize, it is assumed that conditioning on a single variable removes the effect of that variable from correlation coefficients. Because assortment and assortative placement are modeled on the conditional distribution, these relationships are termed *conditional associations* and the parameters, or d_{ij}'s, that index the extent of direct matching between variables are referred to as *conditional parameters* or *conditional paths*. Van Eerdewegh's [12] parameterization uses the Greek δ to denote these relationships. I deliberately choose the symbol d because conditional associations turn out to be a specific case of the general formulation given by Van Eerdewegh. I also follow Van Eerdewegh [12] by indicating conditional pathways with a dashed straight line in a path diagram.

Let k and l denote two other variables, $k \in$ A, and $l \in$ A*. Let $E(kl{:}ij)$ be an operator that denotes the effect of d_{ij} on the correlation between variables k and l. From the definition of d_{ij}, $E(ij{:}ij) = d_{ij}$. $E(kl{:}ij)$ is not necessarily the total expected correlation betwen k and l from the sort/merge, but only that part of the correlation between k and l induced by parameter d_{ij}. Removing the effects of d_{ij} by conditioning on i gives: $0 = E(kl{:}ij) - r(ki)E(li{:}ij)$. Here $r(ki)$ is the correlation between k and i. It has already been derived by using the GLM. $E(li{:}ij)$ may be derived by conditioning on j: $0 = E(il{:}ij) - d_{ij}r(lj)$, so that $E(kl{:}ij) = r(ki)d_{ij}r(jl)$.

This procedure makes the implicit assumption that conditioning reduces the effect to 0, or, in other words, the conditional association is the only reason that two sets are correlated. It may be possible to develop other models where conditioning generates other values of the partial correlation.

To generalize to the case of more than one conditional association, it is assumed that the effect of each d parameter is additive. Thus,

$$r(kl) = \sum_{i} \sum_{j} r(ki) d_{ij} r(jl) \ ,$$

or in matrix form,

$$R_{KK^*} = R_{KM} D R_{M^*K^*} \ , \tag{4}$$

where K is a subset of A and K* a subset of A*. Equation (4) accomplishes the goal by letting K = A and K* = A*. A and A* are Z variables; R_{AM} and $R_{A^*M^*}$ have been calculated by the GLM procedure; and D is a parameter matrix. Consequently, the effects of assortment and/or assortative placement on the Z variables in the two sets can be calculated.

The unstandardized case may be derived by following identical logic for $E(kl{:}ij)$ applied to unstandardized covariances. A simpler alternative is to rescale equation (4). Let S denote a diagonal matrix of standard deviations. Then

$$C_{AA^*} = C_{AM} S_M^{-1} D S_{M^*}^{-1} C_{M^*A^*} \ . \tag{5}$$

Note that D is not rescaled into a covariance matrix. The operator in this case is $S_M^{-1} D S_{M^*}^{-1}$.

Y variables that are causal consequences of A and A* may be derived using the structural equation form given in equation (3): $Y = W_{YA}A + W_{YA^*}A^* + W_{YU}U$, where U denotes all other causal influences not in A and A*. The covariances may be calculated by equation (2) or (3) using C_{AA^*} from equation (5). Note that when variances of A, A*, and U are fixed, conditional associations may change the variance of Y. Consequently, models that begin with standardized variances may not maintain that property under assortment and assortative placement.

AN ALGORITHM TO DERIVE PREDICTED COVARIANCES

The derivations above suggest a succinct, recursive algorithm for deriving predicted covariances in linear models in human genetics. The steps are—(1) *Solve within a set:* Block the model by identifying the two sets joined by conditional associations. Use the GLM to derive predicted covariances for all the variables within each set. (2) *Solve between sets:* Use equation (4) or (5) to derive covariances between the two sets.

These two steps may be continuously repeated until all covariances among Z variables induced by conditional paths are derived. Any remaining structural

equations will be of the form given in equation (1) or (3), so the GLM may be applied to account for additional covariances.

Note that this algorithm uses only two general equations to derive covariances: the general GLM equation (2) including its corollary (3), and the conditional association equation, (4) or (5).

The term "block" in step (1) is a technical term that refers to a series of structural equations that are a subset of the whole model but are solved for simultaneously. The term has not been formally applied to linear modeling in human genetics. When there is more than a single pair of sets joined by conditional associations, correct blocking of a model may be necessary to derive the correct covariances. However, the correct method of blocking a model is more difficult to grasp in the abstract than it is to apply to a concrete model. Van Eerdewegh [12] suggested that temporal sequence dictates the order in which matrices of delta parameters are applied to a model. The recursive nature of the current algorithm also implies that temporal order is a valid criterion for blocking. In order to apply step (2) to sets {A} and {B}, the effects of conditional parameters within sets {A} and {B} must be derived, and events within a set must occur before pairing between sets. Thus, temporal order is used here as a criterion for blocking, although correct covariances may be derived in some models irrespective of temporal sequence. Blocking and the algorithm are illustrated by examples in the next section.

EXAMPLES OF MULTIVARIATE ASSORTMENT AND ASSORTATIVE PLACEMENT

As a first example, consider multivariate phenotypic assortment in a 3-generation pedigree. Assume that variables are unstandardized and that vertical environmental transmission takes place between phenotype of parent and environment of offspring. In temporal sequence, assortment occurs first in the grandparental generation. The sets are {paternal grandfather or pgf} and {pgm}. Within a set, say {pgf}, the vectors are: $Y = P_{pgf}$ and $Z = (G_{pgf}'|E_{pgf}')'$. Placing it into the form of equation (3) gives the structural equations: $G_{pgf} = IG_{pgf}$, $E_{pgf} = IE_{pgf}$, and $P_{pgf} = IG_{pgf} + IE_{pgf}$, which generates the following covariance matrices:

$$C_{GpgfPpgf} = C_{GP} = C_{GG} + C_{GE} ,$$

$$C_{EpgfPpgf} = C_{EP} = C_{EG} + C_{EE} ,$$

$$C_{PpgfPpgf} = C_{PP} = C_{GP} + C_{EP} .$$

P_{pgf} may now be recoded as a Z variable because its covariance matrices with G_{pgf} and E_{pgf} are derived. To model assortment, the vectors are $A = (G_{pgf}'|E_{pgf}'|P_{pgf}')'$ and $M = P_{pgf}$. There are analogous vectors A* and M* for {pgm}. Let $T = SDS$, where $S = \text{diag}(C_{PP})^{-1/2}$. Application of step (2) gives

$$\begin{array}{c} \\ G_{pgf} \\ E_{pgf} \\ P_{pgf} \end{array} \begin{pmatrix} G_{pgm}' & E_{pgm}' & P_{pgm}' \\ C_{GP}TC_{PG} & C_{GP}TC_{PE} & C_{GP}TC_{PP} \\ C_{EP}TC_{PG} & C_{EP}TC_{PE} & C_{EP}TC_{PP} \\ C_{PP}TC_{PG} & C_{PP}TC_{PE} & C_{PP}TC_{PP} \end{pmatrix} .$$

Assortment in the parental generation must now be accounted for, so step (1) is reapplied. The relevant blocks are {pgf, pgm, father (f)} and {mgf, mgm, and mother (m)}. Arbitrarily, select the paternal block to solve for within set covariances; they are assumed to be the same as those within the maternal set. Let t denote a matrix of coefficients of environmental transmission from parent to offspring, and let U denote residuals with subscripts indicating the variates. The structural equations are:

$$E_f = tP_{pgf} + tP_{pgm} + U_{Ef} ,$$

$$G_f = 1/2\ G_{pgf} + 1/2\ G_{pgm} + U_{Gf} ,$$

$$P_f = G_f + E_f = G_f + tP_{pgf} + tP_{pgm} + U_{Ef} ,$$

which give, say, $C_{GfPpgf} = 1/2\ (C_{GP} + C_{GpgmPpgf})$ and $C_{PfPpgf} = C_{GfPpgf} + t(C_{PP} + C_{PpgmPpgf})$. To apply step (2) to the parental generation, first define the relevant vectors. For the paternal set, vector A is the concatenation of the following vectors: G_{pgf}, E_{pgf}, P_{pgf}, G_{pgm}, E_{pgm}, P_{pgm}, G_f, U_{Gf}, E_f, U_{Ef}, and P_f. $M = P_f$. Similar vectors A* and M* may be constructed for the maternal set. Only three relevant matrix blocks are given here:

$$C_{GfPmgm} = C_{GP}TC_{PmPmgm} ,$$

$$C_{PfPmgm} = C_{PP}TC_{PmPmgm} ,$$

$$C_{PpgfPmgm} = C_{PpgfPf}TC_{PmPmgm} .$$

It is easily verified that marital covariances in the parental generation equal those in the grandparents.

Having accounted for assortment in the parental generation, remaining structural equations (i.e., those for grandchild) are all of the GLM form. Covariances between grandchild and parent are assumed to be the same as those between father and pgf. What must be derived are covariances between grandchild and grandparent. The structural equation for grandchild (gc) phenotype is $P_{gc} = 1/2\ G_f + 1/2\ G_m + tP_f + tP_m + U_{Ggc} + U_{Egc}$. Application of equation (3) gives $C_{PgcPpgf} = 1/2\ (C_{GfPpgf} + C_{GmPpgf}) + t(C_{PfPpgf} + C_{PmPpgf})$. All covariance matrices on the right-hand side have already been derived above, so the expected covariance matrix has been calculated.

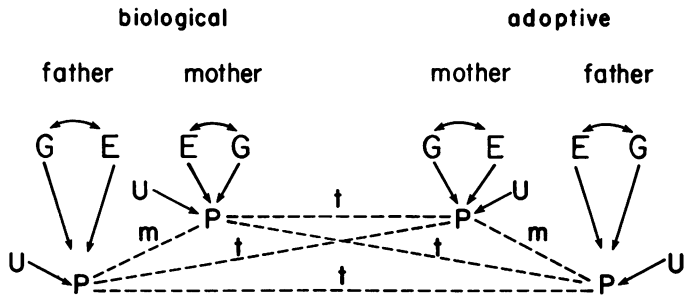

FIG. 1.—A model for assortative placement based on parental phenotypes in the presence of phenotypic mate assortment. G = genotype; E = (systematic) environment; U = unique environment (residual); P = phenotype. *Dashed lines* represent conditional pathways with m denoting phenotypic assortment and t denoting assortative placement based on parental phenotype.

A second example is the solution for combined assortment and assortative placement in an adoption design. It is depicted in figure 1. Figure 1 is schematic in the sense that it is meant to represent a multivariate problem. Hence, the variates G, E, etc., signify vectors and not single variables [10]. The dashed line labeled m denotes a matrix of conditional parameters that models multivariate phenotypic assortative mating. The dashed lines labeled t denote the assortative placement, assumed to take place on the basis of parental phenotypes. For simplicity, it is assumed that t is equal among all four parents. It is also assumed that variables are standardized.

In notation, the Gs and Es constitute the Z vector. In temporal order, mating occurs before assortative placement. Thus, application of step (1) identifies the sets as {husband} and {wife}. Because of symmetry in figure 1, it is immaterial whether they are the biological or adoptive parents. Let h be a matrix of weights (path coefficients) from the vector of phenotypes to the vector of genotypes; e, a weight matrix from the phenotype to the environment vector; and u, a weight matrix from phenotypes to residuals. The structural equation within each set is $P = hG + eE + uU$, giving correlation matrices

$$R_{GP} = R_{GG}h' + R_{GE}e' ,$$

$$R_{EP} = R_{EG}h' + R_{EE}e' ,$$

$$R_{UP} = uu' ,$$

and

$$R_{PP} = hR_{GP} + eR_{EP} + uu' .$$

Vector A may be partitioned as $(G'|E'|U')'$ for one parent with an analogous vector A* for the other parent. Application of equation (4) to these partitions gives the relevant covariances between mates. For example, the correlations

between G of father and P of mother is $R_{GfPm} = R_{GP}mR_{pp}$, and the matrix of marital correlations is $R_{PfPm} = R_{pp}mR_{pp}$.

The algorithm must now be repeated to account for assortative placement. The two sets are {biological father, biological mother} and {adoptive father, adoptive mother}. Correlations within each set have been derived above. A for the biological parents, say, may be partitioned as $(G_f'|E_f'|G_m'|E_m'|P_f'|P_m')'$. An analogous partitioning may be constructed for A* in adoptive parents. $M = (P_f'|P_m')'$, and given symmetrix t, D in equation (4) takes the form

$$\begin{pmatrix} t & t \\ t & t \end{pmatrix}.$$

We do not show the whole partitioned matrix from the application of equation (4) to these definitions of A, A*, and D. Instead, one illustrative matrix block is presented, that between genotypes of biological and adoptive parents: $R_{GbGa} = (R_{GP} + R_{GfPm})t(R_{PG} + R_{PmGf})$. Correlations between adoptive and biological parents are functions of R_{GfPm} and its transpose, which, in turn, are functions of assortment parameters.

If the model were extended to include genotypic, environmental, and phenotypic vectors of adopted offspring, then equation (1) could be applied using the new C_{ZZ} matrix just derived.

NONRECURSION

The algorithm also functions with nonrecursive models, although the application is more intricate than it was for the recursive examples given above. In general, nonrecursive models should be translated into recursive form using antecedent variables, as Carey [17] did in order to model sibling imitation. For example, suppose a marital correlation is a function of two processes, assortment and subsequent imitation. For a single phenotype, structural equations for the latter are $P_f = G_f + E_f + aP_m$, $P_m = G_m + E_m + aP_f$. Recall that the definition of a set includes all variables that occur *before* assortment. One cannot imitate a spouse until an actual pairing has taken place. Thus, assortment takes place on *antecedent phenotypes* of father and mother, say $AP_f = G_f + E_f$ and $AP_m = G_m + E_m$, not on P_f and P_m as defined in the structural equations. Step (2) in the algorithm is applied to the equations for the antecedent phenotypes. The GLM equation (2) can then be applied directly to the structural equations for P_f and P_m. Equation (3) cannot be used without some algebraic reworking because it applies only to recursive models.

DISCUSSION

Starting from first principles and with some explicit assumptions about the form of models, covariance matrices for linear models in human genetics take the form of only two general equations. In the examples, covariance matrices were solved for in a piecemeal fashion using the GLM equation (3) because this

procedure is easy to understand. Using equation (2), however, provides an efficient form for computer algorithms. Instead of tedious algebra or tracing of paths, one can index variables by number, construct the appropriate parameter matrices, and apply equation (2). When variables are ordered by temporal sequence and the Z vector is continuously augmented with each pass through the algorithm, the solution of equation (2) requires only two matrix multiplications. (It is assumed that nonrecursive models have been translated into a recursive form.) The procedure can generate large matrices, but matrix size is offset by the advantage of automatic computation. This is of great importance as models become altered in the course of development. Instead of recomputing expected covariances with the addition of a new path, the appropriate element is simply added to a weight matrix. Also, this procedure is less susceptible to error compared to ad hoc computer programs that perform many matrix manipulations.

Two potential limitations of the present formulation are the application of constraints and the occasional need to block a model. The generic problem with constraints and blocking is that both are model dependent, so it is difficult to make general statements about them. The most frequent constraint in path models of familial resemblance derives from the assumption that variances of genotype and environment, as well as the correlation between the two, are at equilibrium across generations. To solve models involving these assumptions some researchers [1–6] have explicitly derived one parameter in terms of other parameters of the model. In contrast, Heath and Eaves [18] have used constrained optimization. Given the different methods of handling constraints, the application of this algorithm may require ad hoc algebra to apply constraints. This does not appear to be a limitation of this general formulation as much as it is a recognition that constraints themselves are model dependent.

One consequence of multivariate assortment as modeled herein is that the correct multivariate model for assortment must be specified in order to derive correct covariances for the univariate model [12]. The same holds for assortative placement. This is easily verified by deriving the predicted covariances for figure 1 and comparing them to the univariate predictions.

There are several similarities between the present formulation of conditional associations and other attempts to model assortment and assortative placement. Equations (4) and (5) are specific multivariate extensions of Cloninger's [11] formula for copath relationships [his equation (15), not his tracing rule]. The copath in this case is identical to a conditional parameter.

The present formulation differs from Cloninger's in several respects. First, it is assumed here that a conditional path cannot join two variables that are already correlated. Cloninger permits this. Second, conditional associations may change variances of variables further down in the model even when these variables have been previously standardized. Cloninger restricts these variances to remain standardized after a copath is modeled (equation (16) in [11]). Third, the concept of blocking (or simultaneous vs. sequential application of copaths) was not fully explicated by Cloninger. Fourth, he did not discuss nonrecursion.

Van Eerdewegh [12] extended path analysis through the principles of "statistical selection," a term that refers to the formula originally developed by Pearson [19] to predict response to selection and later shown to predict the effects of changes in means, variances, and/or covariances in linear systems [20]. Although I start with different assumptions, equation (4) or (5) can be reworked into Van Eerdewegh's equation (II.5) for delta paths. The present formulation assumes that conditional associations are the only reason that two sets are correlated. Van Eerdewegh's derivations do not require this assumption and consequently are more general. Here, I present a computationally efficient algorithm that joins the GLM with multivariate assortment and allows for nonrecursive models. Van Eerdewegh's interest was in developing the concept of statistical selection and applying it to human genetic problems, so he did not directly address these topics. While the assumption of uncorrelated antecedents limits the mathematical generality of the conditional association approach, it should not impose practical restrictions for many models of assortment or assortative placement; that is, the assumption merely requires that the only reason that the variables of one person are correlated with the variables of his or her mate is the pairing up or matching of mates that occurs in assortment.

ACKNOWLEDGMENTS

I thank Paul Van Eerdewegh for pointing out the identity between the present formulation and his equation (II.5) and thank him, Matt McGue, and two anonymous referees for their comments on an earlier draft. I also thank Rebecca Miles for editorial assistance.

REFERENCES

1. Carey G, Rice J: Genetics and personality temperament: simplicity or complexity? *Behav Genet* 13:43–64, 1983
2. Cloninger CR, Rice J, Reich T: Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance. *Am J Hum Genet* 31:176–198, 1979
3. Eaves LJ, Last KA, Young PA, Martin NG: Model-fitting approaches to the analysis of human behavior. *Heredity* 41:249–320, 1978
4. Fulker DW, DeFries JC: Genetic and environmental transmission in the Colorado Adoption Project: path analysis. *Br J Math Stat Psychol* 36:175–188, 1983
5. Loehlin JC: Heredity-environment analyses of Jencks' IQ correlations. *Behav Genet* 8:415–436, 1978
6. Rao DC, Morton NE, Yee S: Resolution of cultural and biological inheritance by path analysis. *Am J Hum Genet* 28:228–242, 1976
7. Rice J, Cloninger CR, Reich T: Multifactorial inheritance with cultural transmission and assortative mating. I. Description and basic properties of the unitary models. *Am J Hum Genet* 30:618–643, 1978
8. Martin NG, Eaves LJ: The genetical analysis of covariance structure. *Heredity* 38:79–95, 1977
9. McGue M, Rao DC, Reich T, Laskarzewski P, Glueck CG, Russell JM: The Cincinnati Lipid Research Clinic family study: bivariate path analysis of lipoprotein concentrations. *Genet Res* 41:117–135, 1983
10. Vogler GP: Multivariate path analysis of familial resemblance. *Genet Epidemiol* 2:35–54, 1985

11. CLONINGER CR: Interpretation of intrinsic and extrinsic structural relations by path analysis: theory and applications to assortative mating. *Genet Res* 36:133–145, 1980
12. VAN EERDEWEGH P: Statistical selection in multivariate systems with applications in quantitative genetics. Ph.D. thesis, St. Louis, Mo., Washington University, 1982
13. GOLDBERGER AS: Structural equation methods in the social sciences. *Econometrica* 40:979–1001, 1972
14. JORESKOG KG, SORBOM D: *LISREL VI: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood.* Chicago, International Educational Services, 1984
15. GOLDBERGER AS, DUNCAN OD: *Structural Equation Models in the Social Sciences.* New York, Seminar Press, 1973
16. MCARDLE JJ, MCDONALD RP: Some algebraic properties of the reticular action model for moment structures. *Br J Math Stat Psychol* 37:234–251, 1984
17. CAREY G: Sibling imitation and contrast effects. *Behav Genet* 16:319–341, 1986
18. HEATH AC, EAVES LJ: Resolving the effects of phenotype and social background on mate selection. *Behav Genet* 15:15–30, 1985
19. PEARSON K: I. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philos Trans R Soc London A* 200:1–66, 1902
20. AITKEN AC: Note on selection from a multivariate normal population. *Proc Edin Math Soc B* 4:106–110, 1934