

## Cholesky Problems

Gregory Carey  
Department of Psychology and  
Institute for Behavioral Genetics  
University of Colorado  
Boulder CO 80309-0345

Email: [gregory.carey@colorado.edu](mailto:gregory.carey@colorado.edu)  
Phone: 303-492-1658  
Fax: 303-492-2967

Comments appreciated.  
Feel free to cite without permission.

### Abstract

Behavioral geneticists commonly parameterize a genetic or environmental covariance matrix as the product of a nonsingular, lower diagonal matrix postmultiplied by its transpose—a technique commonly referred to as “fitting a Cholesky.” Here, simulations demonstrate that this procedure: (1) may not produce likelihood ratio test statistics that are distributed as a  $\chi^2$ ; or (2) if the distribution of the test statistic appears to be  $\chi^2$ , then the degrees of freedom (df) are not always the difference between the number of parameters in the general model less the number of parameters in the constrained model. It is hypothesized that the problem is related to the fact that the Cholesky parameterization requires that the covariance matrix formed by its product be positive definite. Even though a population covariance matrix may be positive definite, the combination of sampling error and the derived—as opposed to directly observed—nature of some matrices in behavioral genetics allow matrices that are not positive definite. Hence, fitting a Cholesky constrains the area of search and may compromise maximum likelihood theory. Until the reason for this phenomenon is understood and a satisfactory solution is developed, the Cholesky parameterization should be used with caution. An alternate strategy of fitting a lower diagonal matrix to data that avoids the Cholesky problem is proposed.

## Introduction

A common method for modeling genetic and environmental covariance matrices for genetically informative data is to iterate on the elements of a lower diagonal matrix and then obtain the desired covariance matrix by postmultiplying the lower diagonal matrix by its transpose. This procedure is often referred to as “fitting a Cholesky.” This note exposes a potential problem with this approach—namely, in some cases, likelihood ratio test statistics may follow a  $\chi^2$  distribution and the degrees of freedom may not equal the number of free parameters in a general model less the number of free parameters in a nested, constrained model.

## An Illustration of the Problem

I illustrate the problem with a set of 10,000 simulated twin data sets. (Specific details about these simulations and those that follow are given in Methods at the end of this paper). Each data set consisted of 100 pairs of identical and 100 pairs of fraternal twins using only one phenotype with a heritability ( $a^2$ ) of .50, no common environment ( $c^2$ ), and unique environmentability ( $e^2$ ) of .50. For each set of twin data, I computed the intraclass covariance matrix and fitted the same model to the data using two different numerical parameterizations. The first parameterization iterated directly on  $a^2$ ,  $c^2$ , and  $e$ , obtaining  $e^2$  as  $e * e$ . Hence, there are no boundary constraints except for the one implied by iteration on  $e$ .<sup>1</sup> The second parameterization iterated on the elements of the Cholesky, which in this case are simply  $a$ ,  $c$ , and  $e$ , and obtained the variance components as the squares of these estimates.

Table 1 shows the distribution of the  $\chi^2$  goodness-of-fit statistics for the first and second parameterizations according to the percentage of simulations that exceeded the .20, .10, .05, and .01 levels of significance. (Note that the  $\chi^2$  has one degree of freedom

Table 1. Percentage of simulated twin data sets that exceed the critical value for a $p$ level: One phenotype. N equals the number of simulated data sets for a particular condition.				
	Unconstrained	Cholesky, ignoring boundary conditions	Cholesky, fixing parameters to bounds.	
	df = 1	df = 1	Number of parameters fixed:	
			0 (df = 1)	1 (df = 2)
$p$ value	N = 10,000	N = 10,000	N = 4799	N = 5201
.20	20.4	33.1	20.2	21.6
.10	10.4	19.3	10.7	10.6
.05	5.1	10.5	5.1	5.4
.01	1.1	2.5	1.0	1.0

<sup>1</sup> Iteration on  $e$  instead of  $e^2$  improved the conditioning of the model to allow for a much greater proportion of convergences in the simulations.

because models were fitted to the intraclass and not the interclass covariance matrixes.) For the unconstrained parameterization, the distribution is clearly  $\chi^2$  with 1 df. For the Cholesky parameterization, however, what would be the general model is rejected almost twice as often as the unconstrained model.

The reason is easily seen if we consider for the moment only estimates of  $c^2$  (unconstrained) and  $c$  (constrained Cholesky parameterization). Because the population value for  $c^2$  is 0, the unconstrained estimate of  $c^2$  had a mean of 0 with half of the estimates being lower than 0. Iteration on  $c$ , however, cannot give a negative estimate of  $c^2$ , so the value of  $c$  becomes so small that it is effectively fixed at a number very close to 0. Depending upon minimization software (and how that software is used, of course), this fact may or may be brought to the attention of the user. When these boundary conditions are encountered for the single parameter  $c$ , then that  $c$  can be treated as a fixed parameter.

The two rightmost columns of Table 1 illustrate this principle. Here, the 10,000 Cholesky results are divided into two groups, those in which a parameter converged to a bound (arbitrarily defined as an absolute value of the parameter less than .001) and those with no boundary condition. When there is no boundary condition, then the test statistic fits the distribution of a  $\chi^2$  with 1 df. When boundary constraints are met, the test statistic is very close to  $\chi^2$  with 2 df. The fixed parameter was  $c$  in 97.4% of the cases and  $a$  in the remaining 2.6%. In these cases, the numerical estimates of  $c$  or  $a$  ranged from  $5.7 \times 10^{-4}$  to  $4.6 \times 10^{-6}$ .

**The Problem Gets Complicated**

This illustration suggests a simple *ad hoc* solution—keep track of the number of parameters with final values close to 0 and adjust the degrees of freedom accordingly. Unfortunately, this strategy may be appropriate only for the analysis of a single phenotype.

Table 2 presents the summary of simulations that generated two phenotypes on

Table 2. Percentage of simulated twin data sets that exceed the critical value for a $p$ level: Two phenotypes. N equals the number of simulated data sets for a particular condition.						
	Unconstrained	Cholesky, ignoring boundary conditions	Cholesky, fixing parameters to bounds. Number of parameters fixed:			
	df = 3	df = 3	0 (df = 3)	1 (df = 4)	2 (df = 5)	3 (df = 6)
$p$ value	N = 29,986	N = 29,986	N = 1889	N = 16938	N = 6658	N = 4462
.20	19.9	42.3	20.2	24.4	19.8	20.0
.10	10.0	25.4	11.0	13.2	10.0	10.1
.05	5.1	15.1	5.2	7.0	5.4	5.2
.01	1.0	4.2	1.2	1.4	1.0	0.9

100 MZ and 100 DZ pairs. Again, two different parameterizations of the same model were fitted to the data. Both iterated on the Cholesky for the unique environmental matrix. The first parameterization, however, iterated on the unbounded elements in the genetic covariance matrix and the common environmental covariance matrix. The second parameterization iterated on the elements of the Cholesky factors for the genetic and for the common environmental covariance matrices.

The unconstrained solution provides a test statistic that follows a  $\chi^2$  with 3 df, its theoretical distribution. The Cholesky solution, however, rejects the general model more often than expected when the degrees of freedom are set to their nominal value of 3.

The four right-hand columns of the table subdivide the Cholesky results according to the number of boundary conditions encountered (again defined as a parameter with an absolute value less than .001). Note that boundary conditions were more often encountered than not encountered—only 6% of the Cholesky solutions did not encounter a boundary constraint. If this reflects reality, then the problem discussed in this paper may be the norm and not the exception for multivariate situations<sup>2</sup>.

When there are no boundary conditions, the test statistic appears to conform to the expected distribution of  $\chi^2$  with 3 df. Similarly, when two or three parameters are fixed, the test statistic is very similar to the  $\chi^2$  statistics with respectively 5 and 6 df. A problem, however, emerges from that subset where one and only one parameter was set to 0. Here, the distribution is clearly not  $\chi^2$  with 4 df, although the percentage of rejected models at the .05 and .01 levels does not differ markedly from their expected values.

The difficulty here is not in the threshold set for fixing a parameter to a bound. Table 3 gives the percentage of simulations that exceed a significance level as a function of using different criteria for setting a parameter to a bound. In all cases, the percentages exceed the nominal significance level.

Table 3. Effect of different criteria for setting a parameter to a bound of 0. Only those cases in which one parameter was fixed at a bound.				
	Criterion:			
	$ x  < .001$	$ x  < .005$	$ x  < .01$	$ x  < .05$
<i>p</i> level	N = 16,938	N = 16,772	N = 16,451	N = 14,127
.20	24.44	24.44	24.41	23.86
.10	13.20	13.22	13.23	12.73
.05	7.03	7.08	7.08	6.66
.01	1.42	1.42	1.42	1.35

<sup>2</sup> The actual percentage of cases without boundary constraints is influenced by sample size and values of the population covariance matrices, but not greatly so. For example, I repeated these simulations using 1,000 pairs of each zygosity, but still only 15% never encountered a boundary condition. The effect of the population covariance matrix will be discussed later in the paper.

In addition, the problem is more complicated than a simple boundary constraint. It appears that the distribution of the test statistic depends on which parameter hits the boundary constraint. Table 4 gives the results of selecting those cases in which one and only one parameter hit a boundary constraint and then subdividing them by whether that parameter was the lower right-hand element of the genetic Cholesky or the lower right-hand element of the common environmental Cholesky. (In the present simulations, these situations account for over 99.7% of the cases in which only one parameter hit a bound.) If the bounded parameter is in the genetic matrix, then adding a degree of freedom results in a conservative test. But when it is the common environmental parameter, then the result is a liberal test.

Table 4. Percentage of simulations exceeding a critical level when the lower right-hand element of the genetic or common environmental Cholesky hit a bound: Two phenotypes using only simulations where one and only one parameter hit a bound.				
	Parameter with the Boundary Constraint:			
	Genetic Parameter N = 2,271		Common Environment Parameter N = 14,631	
<i>p</i> level	df = 3	df = 4	df = 3	df = 4
.20	28.2	16.5	40.4	25.7
.10	14.8	8.5	23.6	13.9
.05	8.2	4.1	13.8	7.5
.01	1.9	0.7	3.4	1.5

In some cases, the problem appears to be one of the degrees-of-freedom for a  $\chi^2$  distribution and not the possibility that the distribution is something other than  $\chi^2$ . Figure 1 gives a histogram of the function value when the genetic parameter hits a bound and Figure 2, when the common environment parameter hits a bound for all those simulations in Table 4. Plotted with each histogram is the  $\chi^2$  probability density function for the degrees of freedom that best fit the data. For the genetic parameter, these degrees of freedom equal 3.7 and for the environmental parameter they are 4.6.

Figure 1. Observed distribution of the goodness-of-fit  $\chi^2$  statistic when a parameter in the genetic Cholesky encountered a bound, along with the theoretical  $\chi^2$  distribution with the best fitting degrees of freedom.

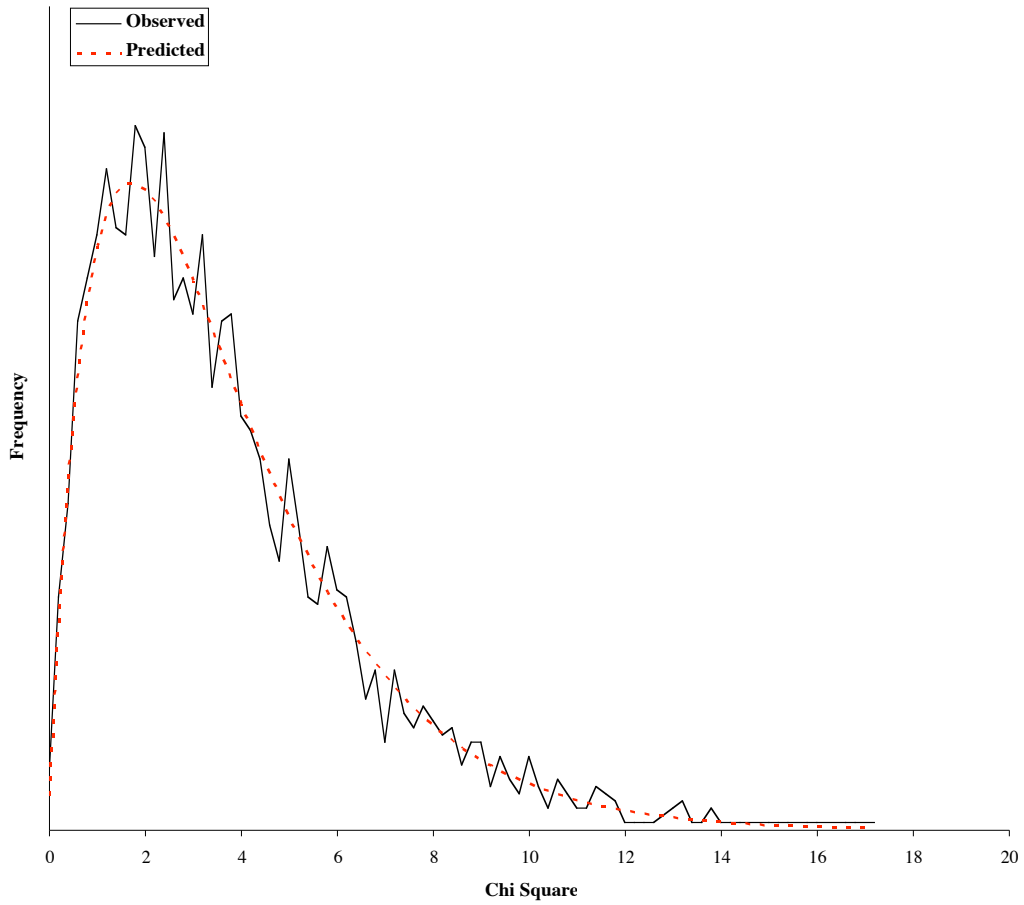
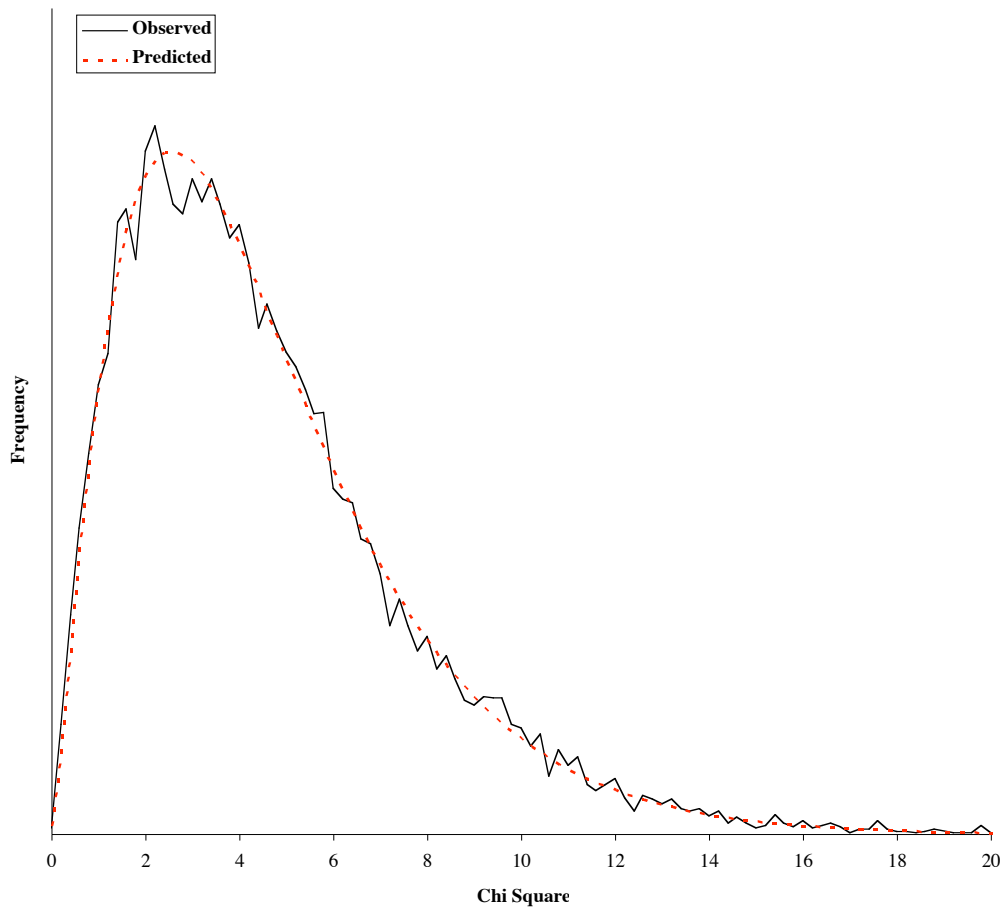
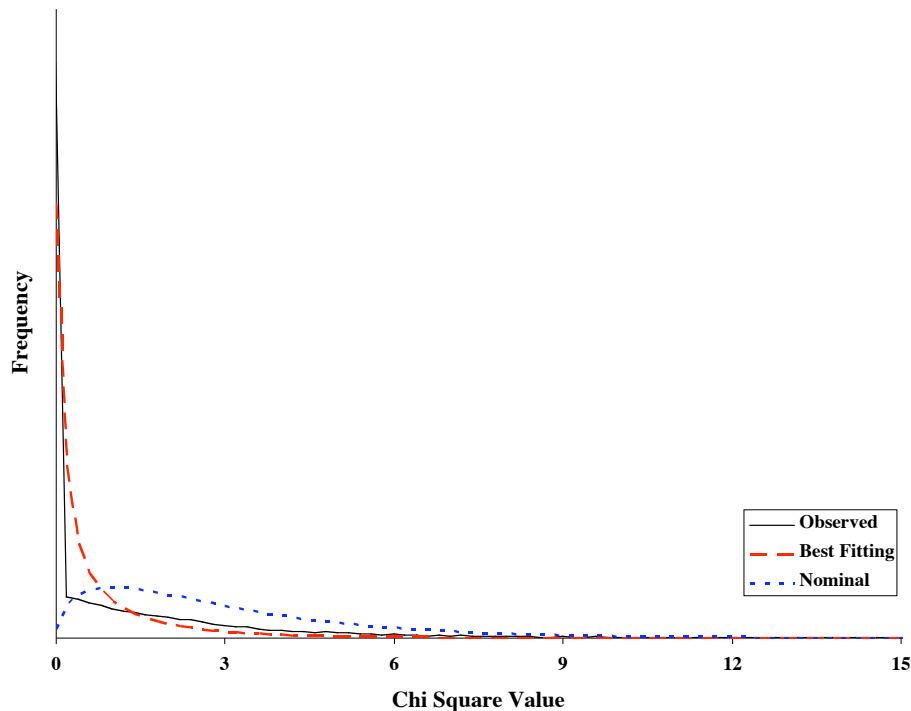


Figure 2. Observed distribution of the goodness-of-fit  $\chi^2$  statistic when a parameter in the common environment Cholesky encountered a bound, along with the theoretical  $\chi^2$  distribution with the best fitting degrees of freedom.)



In other cases, however, the distribution itself is not  $\chi^2$ . Figure 3 gives the distribution of the likelihood ratio  $\chi^2$  that tests whether all three elements of the Cholesky factors for common environmental covariance matrix are 0. In the “business as usual approach,” such a test is regarded as having three df. The line labeled “df: Nominal” gives the theoretical  $\chi^2$  distribution for 3 df, and it is clear that the observed test statistic is nowhere close to matching this distribution. The line labeled “df: Best Fitting” plots the  $\chi^2$  distribution with the best fitting degrees of freedom (which were 0.64 in these simulations). This line does not fit the observed data well.

Figure 3. Observed, theoretical, and best-fitting  $\chi^2$  distributions for the likelihood ratio statistic testing whether all elements in the Cholesky common environment matrix are 0.



Frankly, I am unable to offer a formal mathematical proof of why all these phenomena occur. Some post hoc exploration of the results of the simulations—and many other simulations not reported here—suggest that the problem is intimately related to the eigenvalues of the unconstrained estimates of the covariance matrices. Table 5 gives the number of boundary conditions encountered as a function of the number of positive eigenvalues for the unconstrained genetic covariance matrix (matrix **A**) and the unconstrained common environmental covariance matrix (matrix **C**).

When both matrix **A** and **C** are positive definite, then only 2 of 1,817 (0.1%) of the simulated data sets encountered a boundary constraint under the Cholesky parameterization. Given that the limit for setting a parameter to a bound is arbitrary and may include some cases in which the parameter value is truly very small, these 2 cases may represent false positives. Of the 28,169 data sets in which neither **A** and/or **C** was positive definite, only 74 (0.3%) were free of boundary constraints. Additionally, the number of positive eigenvalues is negatively correlated with the average number of parameters encountering a bound. When 3 or the four eigenvalues are positives, then the number of parameters hitting a bound averaged between 1 and 1.1. When only two of the four eigenvalues were positive, then somewhere between 2 and 3 parameters encountered a boundary constraint.

Table 5. Relationship between the signs of the eigenvalues of the unconstrained genetic (A) and common environment (C) matrices and boundary conditions on elements of their Cholesky matrices: Two phenotypes.								
Number of Positive Eigenvalues in Matrix		Number of Parameters Encountering a Boundary Condition:						Average Number of Bounded Parameters:
A	C	N	0	1	2	3	4	
2	2	1817	1815	2	0	0	0	0.0
2	1	15451	74	14641	722	10	4	1.0
2	0	4465	0	21	3	4420	21	3.0
1	2	2416	0	2223	193	0	0	1.1
1	1	5817	0	51	5740	12	14	2.0
1	0	0						
0	2	20	0	0	0	20	0	3.0
0	1	0						
0	0	0						

Not only do the eigenvalues of the unconstrained matrix predict the number of boundary constraints, but they also predict the specific parameters that enter into those constraints. Given that unconstrained **A** was positive definite, then the probability of fixing an element in the Cholesky for **A** was only .034. Given that **A** was negative definite, however, the probability of fixing one of its Cholesky elements was .999. Given that unconstrained **C** was positive definite, the probability of fixing an element in its Cholesky was .046. When **C** was negative definite, however, the probability was .995.

This patterning holds for the more complicated situations in Table 5. For example, when both eigenvalues of matrix **C** were negative, then 99% of the time three parameters encountered boundary conditions. In virtually every case, these three parameters were the three elements in the Cholesky factors of matrix **C**.

The final observation about the eigenvalues is that the status of unconstrained matrix **A** is not independent of the status of unconstrained matrix **C**. If we categorize each matrix into the binary class of positive definite versus negative definite, then  $\chi^2 = .28$ . Hence, when matrix **A** is positive definite—which, given the parameter values that generated these simulations, is the typical case—then there is a small tendency for matrix **C** to have at least one negative eigenvalue.

Excepts for the points noted immediately below, simulations using three, four, and five phenotypes gave similar types of results to those presented above and are not reported in detail here. In these simulations, the heritability for each phenotype was the same (0.5), the genetic correlation among any two phenotypes was the same (0.6), and the unique environmental correlation among any two phenotypes was the same (0.2). All common environmental effects were set to 0.

The major effect of adding more phenotypes is to increase both the intensity and the complexity of the problem. The percentage of Cholesky models rejected at the .05 level by the goodness-of-fit  $\chi^2$  was 21.2% (three phenotypes), 30.3% (four phenotypes),

and 47.0% (five phenotypes). Similarly, the percentage of simulations in which the search entered acceptable space (i.e., all eigenvalues for both the unconstrained genetic and the unconstrained common environment matrix were positive) dropped precipitously—0.5% (three phenotypes), 0.02% (four phenotypes), and 0% (five phenotypes).

### Discussion

What factor(s) are responsible for these problems and how can they be remedied? I lack the mathematical and statistical background to provide firm answers to these questions. But on the basis of the simulations and—quite frankly—some guesswork, I offer an explanation that is probably not correct but, I hope, will assist others to develop more definitive and appropriate answers to this question. Before doing so, a brief digression about Cholesky factorization will be helpful.

Textbooks in linear algebra define Cholesky factorization as a process that begins with a positive definite, symmetric matrix, say  $\mathbf{X}$ , and then factorizes  $\mathbf{X}$  into a nonsingular, lower diagonal matrix,  $\mathbf{W}$ , such that  $\mathbf{X} = \mathbf{W}\mathbf{W}^t$ . A corollary is that if  $\mathbf{W}$  is a square, nonsingular, lower diagonal matrix, then the matrix formed by  $\mathbf{W}\mathbf{W}^t$  must be positive definite. The key element here is the word “must.” It is a mathematical necessity that  $\mathbf{X}$  is positive definite as long as each element on the diagonal of  $\mathbf{W}$  is different from 0.

Models in behavioral genetics differ from those in other areas of social science in that they predict covariance matrices (technically, covariance matrices blocks) that are not directly observed but are: (1) mathematical functions of other matrix blocks; and/or (2) a block or subset of an observed correlation matrix. In the simplest case of the second example, imagine a data set consisting of a single phenotype, say IQ, measured on identical twins raised apart. The observed correlation matrix is a two by two matrix in which the first variable is twin 1’s IQ score and the second is twin 2’s IQ score. The correlation between twin 1’s IQ and twin 2’s IQ is that block within this two by two matrix that directly estimates heritability for IQ. Now ask whether it is a mathematical necessity (as opposed to a simple probability) that this observed correlation be positive. We all agree that the answer is a simple probability and not a mathematical necessity, the actual probability being dependent on the population value of heritability and the number of twin pairs in the sample.

Now let us take two mathematical models and fit them to these hypothetical data on twins raised apart. The first model allows for the mathematical possibility—however improbable—that the correlation between identical twins could be negative. This model iterates on a quantity such as the heritability,  $h^2$ , but operates on the entire area of real numbers. Hence, legitimate values of  $h^2$  could be negative as well as positive. The second mathematical model entirely precludes the possibility that sampling error could result in a negative observed correlation. Again, this model iterates on the quantity  $h^2$ , but performs the search in only that area of the real numbers in which  $h^2$  is greater than 0.

My conjecture is that the first of these two mathematical models—the one that permits the unlikely event of a negative correlation—is more appropriate than the second. This model allows all possible parameter values to be elements of the real numbers,

irrespective of sign. Hence, this model accounts for the possibility of sampling error that could give a correlation between identical twins raised apart that is less than 0.

The second model constrains the permissible area of parameter values for heritability to be positive. Although it is appropriate to *logically* constrain estimates of heritability to be elements of the positive real numbers, this constraint eliminates all the mathematical space that permits the heritability to be negative. In effect, it denies the possibility that sampling error could result in a negative twin correlation. Hence, the maximum likelihood surface is constrained and traditional likelihood theory is compromised.

The Cholesky parameterization limits the maximum likelihood search area to the space in which matrices (genetic, common environment, and unique environment) are all positive definite. Hence, instead of searching on a minimum in the multidimensional space consisting of all the real numbers, the search is restricted to a local minimum within the space of positive definite matrices. For a single phenotype, the solution is simple because the constrained search area is equivalent to a boundary condition—fix the parameter to its bound and add a degree of freedom. I cannot, however, explain why such logic does not pertain to the case of more than one phenotype.

One hopes that this explanation is not entirely correct because it portends problems for virtually any method of parameterizing the covariance matrices for latent variables in behavioral genetic analysis that does not permit a search over the whole space of real numbers. It may be the case that many different types of variance components models—from those used here to the ones used in linkage and association studies—could be influenced by this problem. Clearly, we need to explore this important area.

A very different question is whether the Cholesky problems outlined above create insurmountable problems for behavioral genetic analysis. Here, the answer is “yes” as long as we continue to do business as usual. In the simulations described above—and in the theoretical explanation provided herein—Cholesky problems will most likely be encountered with matrices with population values close to 0. Why? Because sampling error is more likely to result in an unconstrained variance estimate that is less than 0 in such a matrix than in a matrix with population values much larger than 0. Given the empirical estimates most often encountered in behavioral genetics, then Cholesky problems are most likely to be found in matrices involving the common environment than with genetic or unique environmental covariance matrices.

The “business as usual” approach fits Cholesky factors to the common environment covariance matrix, then fits a second model in which all elements of the Cholesky are set to 0, and assumes that the nominal degrees of freedom for the likelihood ratio test are  $n(n + 1)/2$  where  $n$  is the number of phenotypes. As the simulations have demonstrated, however, the distribution—if it is even  $\chi^2$  to begin with—is unlikely to have these degrees of freedom, and the best fitting df will be something less than  $n(n + 1)/2$ . If there were five phenotypes, then there 15 free elements in the Cholesky, giving a test with ostensibly 15 df. If, in fact there were only 9.7 df because of the Cholesky constraints, then we lose power to detect some common environmental effects. Champions of the family environment may justifiably look askance at field that claims that there is little common environment when the statistical techniques used to justify that assertion are not fully powered.

Sample size obviates some of the difficulties with the Cholesky, but it does not completely eliminate the problem. In simulations not reported here, increased sample size reduced problems with the genetic covariance matrix, but it had little effect on the common environment matrix. The reason for this is easily illustrated by considering those simulations using only a single phenotype. Here, the only time a Cholesky problem occurs with the genetic covariance matrix is when sampling error gives a negative heritability. When the heritability is moderate—it was 0.5 in the simulations—then increasing the number of twin pairs makes it much less likely that the DZ correlation will, by dumb luck, exceed the MZ correlation. When there is no common environment, however, then half of the time the unconstrained estimate of  $c^2$  will be negative regardless of sample size.

So what should behavioral geneticists do? For the present, we should avoid using Cholesky parameterizations until the whys-and-wherefores of the Cholesky problems are better explicated and understood than could be done in this paper. For certain types of problems, this will provide no great disadvantage.

But what should one do when substantive considerations argue that the form of a matrix should be lower diagonal? Good examples include developmental models in which genetic effects for a time point can influence phenotypes at the same and subsequent time points but not phenotypes at prior times. Here, one can use Gaussian elimination to show that any nonsingular, symmetric matrix,  $\mathbf{X}$ , can be factored so that  $\mathbf{X} = \mathbf{L}\mathbf{D}\mathbf{L}'$  where  $\mathbf{L}$  is a lower diagonal matrix with 1.0 on each diagonal element and  $\mathbf{D}$  is a diagonal matrix. By allowing elements of  $\mathbf{D}$  to be negative, one eliminates the problem that  $\mathbf{X}$  be positive definite.

It can be shown that this parameterization will give the same answer as the unconstrained solution in the general case. There are, however, two problems with this approach, one rather minor but the other substantive. First, numerical searches on this parameterization using arbitrary starting values are not well conditioned, so it often takes a number of different tries to achieve convergence. The substantive problem occurs when the solution has a negative element in  $\mathbf{D}$ . This denotes a negative variance, and how does one interpret a negative variance? The problem goes away if one can set that element in  $\mathbf{D}$ —and, of course, all elements in the corresponding column of  $\mathbf{L}$ —to 0 without a worsening in fit. But this may not always be the case.

It is not clear whether this parameterization is completely free of problems, and more research is needed in this area. Particularly, the effect of a singular  $\mathbf{X}$ —or close to singular  $\mathbf{X}$ —needs to be explored. Still, I hope that this will provide a suitable stopgap until the Cholesky problem can be resolved.

## Methods

Software for the simulations was developed by the author as a series of programs and subroutines written in Fortran 90 and 95 that used algorithms from the NAG (Numerical Algorithms Group, <http://www.nag.com>), LAPACK, and BLAS libraries for random number generation, function minimization, and certain matrix operations. In each simulated data set, genetic values were generated for a twin pair based on simple additive model with a genetic correlation of 1.0 for MZ twins and 0.5 for DZ twins. Here, pseudorandom numbers were generated from a multivariate normal distribution with a mean of 0 and a pre-specified genetic covariance matrix (and, of course, the

relevant genetic correlation for the type of twin pair) giving vectors  $\mathbf{g}_1$  for twin 1 and  $\mathbf{g}_2$ . Two uncorrelated vectors of unique environmental values,  $\mathbf{e}_1$  and  $\mathbf{e}_2$ , were separately generated from a multivariate normal distribution with means of 0 and a pre-specified unique environmental covariance matrix. Phenotypes for a twin were then computed as the sum of the relevant genetic and unique environmental vectors.

Intraclass covariance matrices were calculated for MZ and for DZ twins based on the phenotypic vectors by double entry of the twin pair in reversed order. The function minimized was

$$-2\text{Log}(L) = \sum_{i=mz}^{dz} (N_i(\log \hat{\Sigma}_i - \log S_i) + \text{trace}(\hat{\Sigma}_i^{-1} S_i) - p)$$

where  $N_i$  is the number of twin pairs for a zygosity,  $\hat{\Sigma}_i$  is the predicted and  $S_i$ , the observed covariance matrix, and  $p$  is the number of phenotypes.

The predicted phenotypic covariance matrix block was the sum of matrices  $\mathbf{A}$  (additive genetic covariance matrix),  $\mathbf{C}$  (common environmental covariance matrix) and  $\mathbf{E}$  (unique environment covariance matrix), and the cross-twin covariance was calculated as  $\lambda\mathbf{A} + \mathbf{C}$  where  $\lambda$  equaled 1.0 for MZ and 0.5 for DZ twins. In the unconstrained model, elements of  $\mathbf{A}$  and  $\mathbf{C}$  were all free parameters that were estimated without bounds. In the Cholesky model, elements of Cholesky factors  $\Sigma_A$  and  $\Sigma_C$  were iterated on so that  $\mathbf{A} = \Sigma_A \Sigma_A^t$  and  $\mathbf{C} = \Sigma_C \Sigma_C^t$ . In both the unconstrained and Cholesky models, elements the lower diagonal matrix  $\Sigma_E$  were iterated on so that  $\mathbf{E} = \Sigma_E \Sigma_E^t$ . (Given uncorrelated error variance, matrix  $\mathbf{E}$  must be positive definite and the current parameterization greatly improved the numerical search. No element in matrix  $\Sigma_E$  ever entered into a boundary constraint in the simulations reported here.)

Two sets of starting values were used, one based on the observed covariance matrices and the other from the genetic and unique environmental covariance matrices used to generate the simulated data. Minimization of the function was done using the NAG algorithm e04ucf. Two criteria were used to assess convergence. The first was the NAG return value indicating successful convergence along with a test that the convergence was not obtained because the parameters entered a space where a predicted twin matrix was not positive definite. If, despite using difference start values, the NAG return did not indicate successful convergence but instead suggested numerical problems in the accuracy of the calculated function value or of elements in the Hessian matrix, then first derivatives were calculated. If the norm of the gradient was less than  $10^{-6}$  and no derivative exceeded .0001, then convergence was assumed. Lack of convergence was not a problem. It never occurred using one phenotype and occurred .05% of the time with two phenotypes.

The parameters used for a single phenotype were given in the text. The genetic covariance matrix for the two-phenotype simulations was

$$\begin{bmatrix} .52 & .48 \\ .48 & .52 \end{bmatrix},$$

and the unique environmental covariance matrix

$$\begin{bmatrix} .5125 & .39 \\ .39 & .5125 \end{bmatrix}.$$

Many other covariance matrices were also used, but I chose to present these because they were the ones used to develop and test the program and to perform the initial simulations.

There are circumstances in which the Cholesky problems are minimized but they hardly constitute meaningful conditions for multivariate analysis. One clear example is to have a very large number of twin pairs and diagonal genetic, common environment, and unique environment matrices where the proportion of variance explained by each of the three is one-third. Sampling error that would rarely result in a negative variance in such cases, but then how informative is a multivariate analysis of uncorrelated phenotypes?

The parameter values used for the simulations not reported above were largely in the range of those often encountered in behavioral genetic research—moderate heritabilities, genetic correlations that largely explained phenotype correlations, and little common environment. The substantive conclusions of this paper were robust over this parameter space.

For a single phenotype, 10,000 replicates were generated. I initially used 10,000 replicates for two phenotypes, but they provided too few cases of the rare combinations of parameters that encountered bounds to permit further exploration. Hence, I used 30,000 replicates. Only those simulated data sets in which the unconstrained and the Cholesky models both converged were further analyzed.