Proofs for Algorithm

Dayou Luo

2018/7/25

Inverse of a covariance matrix

4-dimensional index

First, I will introduce a 4-dimensional index to label elements of a $n(n-1)$ square matrix.

This idea comes from the covariance matrix of residuals. In a covariance matrix of residuals, there are $n \times (n-1)$ rows and $n \times (n-1)$ columns. Each item in the matrix stands for a covariance of two residuals, i.e. $cov(e_{(i,j)}, e_{(m,n)})$, where (i,j) subscript illustrates that this residual comes from the interaction of the i^{th} and j^{th} actor. Thus there will be a unique map connecting the row and column index to a 4-dimensional index, (i,j,m,n), where (i,j,m,n) is the subscript of the corresponding residuals. To be more specific, I am going to define a 4-dimension notation,(i,j,p,q), for an arbitrary square matrix with $n \times (n-1)$ rows. For the element on the a^{th} row and b^{th} columns, define (i,j) as the solution of :

$$a = ((i-1) + (j-1) \times (n-1) + (j > i)) \times (i! = j) \quad i > 0,\ j > 0,\ i,j \ are\ integers,$$

And (p,q) as the solution of:

$$b = ((p-1) + (q-1) \times (n-1) + (q > p)) \times (p! = q), \quad p > 0,\ q > 0,\ p,q \ are\ integers,$$

where

$$(j > i) \ = \begin{cases} 1 & when\ j < i, \\ 0 & when\ j \geq i, \end{cases}$$

and

$$(i! = j) = \begin{cases} 1 & when\ i \neq j, \\ 0 & when\ j = i, \end{cases}$$

It is easy to prove that (i,j,p,q) is unique with (a,b) clarified.

An exchangeable matrix

Next, I will introduce the exchangeable property for a matrix. In a covariance matrix of residuals of interactions, arbitrary permutation π, which switches the order of actors, will not change the covariance matrix. That is to say that $cov(e_{(i,j)}, e_{(m,n)}) \ = \ cov(e_{(\pi(i),\pi(j))}, e_{(\pi(m),\pi(n))})$. We can use the 4-dimension notation for this definition: setting A as the covariance matrix, $A_{i,j,m,n} \ = \ A_{\pi(i),\pi(j),\pi(m),\pi(n)}$. Extending this definition to an arbitrary matrix, I define a matrix as an exchangeable matrix if it follows restrictions listed below:

1. A matrix has $n \times (n-1)$ rows and $n \times (n-1)$ columns.
2. With the 4-dimension notation, a matrix is unchanged under any permutation of each dimension. That is to say that: setting the matrix as A, than: $A_{i,j,m,n} \ = \ A_{\pi(i),\pi(j),\pi(m),\pi(n)}$, in which π is an arbitrary permutation for integers within 1 to N and (i,j,k,l) are arbitrary numbers while $i \neq j$ and $k \neq l$.

By definition, it is easy to check that all the covariance matrix of residuals are exchangeable with the exchangeable assumption. Also, an identity matrix is exchangeable.

However, covariance matrices are special exchangeable matrices. Under the general definition, an exchangeable matrix can have at most 7 different values. On the other hand, a covariance matrix only can have 6 unique values because it is symmetric. To be specific, these 7 different values can be written as following. Here i,j,k,l stand for different integers ranging from 1 to N, A is one exchangeable matrix.

$$
\begin{aligned}
&A(i,j,i,j)=\phi_1 \quad A(i,j,j,i)=\phi_2 \quad A(i,j,i,k)=\phi_3\\
&A(i,j,k,j)=\phi_4 \quad A(i,j,k,i)=\phi_5 \quad A(i,j,k,l)=\phi_6\\
&A(i,j,j,k)=\phi_7. \quad If\ A\ is\ symmetric,\ \phi_5=\phi_7.
\end{aligned}
$$

Proof

With the 4-dimensional notation and the definition of an exchangeable matrix, we can work on the proof of the conjecture:if a covariance matrix is invertiable, its inverse an symmetric exchangeable matrix.

Lemma: if A, B are exchangeable matrics, $A \times B$ will also be an exchangeable matrix.\

Proof: Set $C = A \times B$. Consider the element on the on the a^{th} row and b^{th} columns of C, noted as $C_{(a,b)}$. By defination,

$$C_{(a,b)} = \sum_{p=1}^{n\times(n-1)} A_{(a,i)} \times B_{(i,b)}$$

Switching the notation to the 4-time notation. (a, b) will be mapped to (i, j, k, l). (i, j) is equivalent to the a^{th} row of the matrix and (k, l) will represent the b^{th} column of the matrix. With the same mapping rule, $A_{(a,p)}$ will be $A_{(i,j,M(p),N(p))}$, which has the same first two coordinate as C. $B_{(p,b)} = B_{(M(p),N(p),k,l)}$

Thus,

$$
\begin{aligned}
C_{(i,j,k,l)} &= \sum_{p=1}^{n\times(n-1)} A_{(i,j,M(p),N(p))} \times B_{(M(p),N(p),k,l)}\\
&= \sum_{q=1}^{n} \sum_{r\neq q, r\leq n} A_{(i,j,q,r)} \times B_{(q,r,k,l)}\\
&= \sum_{q=1}^{n} \sum_{r\neq q, r\leq n} A_{(i,j,\pi_0(q),\pi_0(r))} \times B_{(\pi_0(q),\pi_0(r),k,l)}, \quad for\ \forall\ permutation\ \pi_0.
\end{aligned}
$$

The second equality holds because $(M(p), N(p))$ goes through every combination of $(i, j),\ i \neq j$ as p grows from 1 to $n \times (n - 1)$. The third equation holds since π only changes the order of summation.

Consider an arbitrary permutation π, changing C to C'. I have that:

$$
\begin{aligned}
C'_{(i,j,k,l)} &= C_{(\pi(i),\pi(j),\pi(k),\pi(l))}\\
&= \sum_{q=1}^{n} \sum_{r\neq q, r\leq n} A_{(\pi(i),\pi(j),q,r)} \times B_{(q,r,\pi(k),\pi(l))}\\
&= \sum_{q=1}^{n} \sum_{r\neq q, r\leq n} A_{(\pi(i),\pi(j),\pi(q),\pi(r))} \times B_{(\pi(q),\pi(r),\pi(k),\pi(l))}\\
&= \sum_{q=1}^{n} \sum_{r\neq q, r\leq n} A_{(i,j,q,r)} \times B_{(q,r,k,l)}, \quad by\ exchangeablity\ of\ A\ and\ B\\
&= C_{(i,j,k,l)}
\end{aligned}
$$

The third equality is a consequence of the matrix multiplication of 4-dimensional index proved above. This shows that: C is exchangeable.

Therefore, C has 7 unique parameters. If we set these parameters as $q_7 = [1, 0, 0, 0, 0, 0, 0, 0]^T$, we can get the 7 unique parameters of the covariance matrix's inverse through the following linear equation:

$$C_7(\phi_{A,7}, n)\phi_{B7} = {q_7}^T$$

Where the first 5 columns of C_7 are

$$\begin{bmatrix}
\phi_1 & \phi_2 & (n-2)\phi_3 & (n-2)\phi_4 & (n-2)\phi_7 \\
\phi_2 & \phi_1 & (n-2)\phi_5 & (n-2)\phi_7 & (n-2)\phi_4 \\
\phi_3 & \phi_7 & \phi_1+(n-3)\phi_3 & \phi_5+(n-3)\phi_6 & \phi_2+(n-3)\phi_7 \\
\phi_4 & \phi_5 & \phi_7+(n-3)\phi_6 & \phi_1+(n-3)\phi_4 & \phi_3+(n-3)\phi_6 \\
\phi_5 & \phi_4 & \phi_2+(n-3)\phi_5 & \phi_3+(n-3)\phi_6 & \phi_1+(n-3)\phi_4 \\
\phi_6 & \phi_6 & \phi_4+\phi_7+(n-4)\phi_6 & \phi_3+\phi_5+(n-4)\phi_6 & \phi_3+\phi_5+(n-4)\phi_6 \\
\phi_7 & \phi_3 & \phi_4+(n-3)\phi_6 & \phi_2+(n-3)\phi_7 & \phi_5+(n-3)\phi_6
\end{bmatrix}$$

The last two rows are:

$$\begin{bmatrix}
(n-2)(n-3)\phi_6 & (n-2)\phi_5 \\
(n-2)(n-3)\phi_6 & (n-2)\phi_3 \\
(n-3)(\phi_4+(n-4)\phi_6+\phi_5) & \phi_4+(n-3)\phi_6 \\
(n-3)(\phi_3+(n-4)\phi_6+\phi_7) & \phi_2+(n-3)\phi_5 \\
(n-3)(\phi_3+(n-4)\phi_6+\phi_7) & \phi_7+(n-3)\phi_6 \\
(n-4)(\phi_3+\phi_7+\phi_4+\phi_5+(n-5)\phi_6)+\phi_1+\phi_2 & \phi_4+\phi_7+(n-4)\phi_6 \\
(n-3)(\phi_4+(n-4)\phi_6+\phi_5) & \phi_1+(n-3)\phi_3
\end{bmatrix}$$

The matrix built with ϕ_B will be the inverse of covariance matrix. It is easy to prove that the inverse matrix is symmetric. Therefore, in the inverse matrix, $\phi_{B5} = \phi_{B7}$. Thus, we proved that if a covariance matrix is invertiable, its inverse is a symmetric exchangeable matrix, which has the same data pattern as the covariance matrix.

As a result, we can use the following function, which is a reduced form of the equation above, to get the inverse matrix.

$$C_6(\phi_{A6}, n)p_{B6} = q_6^T \quad for\ A(\phi, n)\ \in R^{6\times 6}.$$

ϕ_{A6} is the six unique parameters of A and p_{B6} is that of B. q_6 is the six unique parameters of C. $C_6(\phi, n)$ is a 6×6 matrix elaborated in the original paper.

Getting p from the equation with $q_6 = [1, 0, 0, 0, 0, 0, 0]^T$ and ϕ from A, I can build an exchangeable matrix with p_{B6}, noted as B, when $C_6(\phi, n)$ is not singular. b is the inverse we want.

While C_7 is invertible,C_6 will be invertible. But I can't be sure whether the converse proposition is still correct. Probably there is some ϕ making C_7 singular and C_6 nonsingular. In this case, the matrix constructed with the outcomes of C_6 will not be the inverse of covariance matrix. However if we can prove that C_7 is invertible with probability 1, then C_6 is equivalent to C_7. That is probably why we didn't find couter-example with random samples. Unfortunately, I can't prove that C_7 is invertible with probability 1, and I may need some help here.

Faster algorithm for missing data.

When there are missing data, our present algorithm directly inverses the covariance matrix of residuals. This is not efficient: saving the whole matrix will cost huge memory space and the computation will be time-consuming. Consider a relation array with 100 actors and suppose that 1% of the data is missing. The covariance matrix will be a matrix that has approximately 10000 rows and 10000 columns. If a double item takes 8B memory space, the whole matrix will take over 700MB memory space. Inversing such a matrix will be a disaster.

A better method takes advantage of the exchangeable property of the covariance matrix in order to speed up the calculation. Instead of inverting the covariance matrix of existed data, we can inverse the covariance matrix of missed data, which normally has a much smaller size. The new strategy can also take advantage of the spares property of the covariance matrix. Using the same data, the new method only directly inverses a matrix with 100000 items. This change makes the method feasible for the missing data situation when the size of the missed data is not large. If there is more missing data than data existed, the new method will become less efficient. However, with such size of missing data, the research itself will be not convincing. One should include more data or change to another dataset to get a more accurate outcome.

Here is the algorithm:

1.Generate the covariance matrix as if no data is missing. Mark the matrix as A.

2.Generate the inverse of A. Mark it as B.

3.Pick the rows that represent existed data, and switch these rows to the left of the matrix. Pick the columns that represent existed data and move them up. As a result, the up-left part of the covariance matrix is the covariance matrix of all existed data. This step will generate a new matrix. I call it the matrix A'. The left-up part will be marked as A'_1.

4.Do exactly the same operation to B to generate B'. It can be proved that $A' \times B' = I$

5.Partition the B matrix into 4 parts, left-up, left-down, right-up and right-down. There will be as many columns and rows in the left-up matrix as the size of data. Mark these matrics as $B'_1, B'_2, B'_3 \ and \ B'_4$.

6.${A'_1}^{-1} = B'_1 \ - \ B'_2{B'_3}^{-1}B'_4$. It worth noting that $B'_1, B'_2, B'_3 \ and \ B'_4$ are all sparse matrices.

7.Suppose we want to calculate $X{A'_1}^{-1}Y$ where X and Y are two matrices. We have that:

$$\begin{aligned} X{A'_1}^{-1}Y &= X(B'_1 \ - \ B'_2{B'_3}^{-1}B'_4)Y \\ &= XB'_1Y \ - \ (XB'_2{B'_3}^{-1})B'_4Y \end{aligned}$$

Since B'_1, B'_2, B'_4 are sparse matrices, sparse notation can save a lot of memory space and computing time.

I need to prove two lemmas for validating this algorithm.

The first lemma is to prove that: $A' \times B' = I$. Rewrite A' as $A' = P_{i_1,j_1} \times P_{i_2,j_2} \times \ldots P_{i_n,j_n} \times A \times P_{i_n,j_n} \ldots \times P_{i_1,j_1}$ in which $P_{i,j}$ is a permutation matrix with (i,j) =1, (j,i)=1 and (k,k)=1 when $k \neq i$ and $k \neq j$. The other items in $P_{i,j}$ are zero. $P_{i,j}$ will permute the i^{th} row and j^{th} row when multiplied on the left of a matrix. A property of $P_{i,j}$ is $P_{i,j} \times P_{i,j} = I$. Thus

$$\begin{aligned} A' \times B' &= P_{i_1,j_1} \times P_{i_2,j_2} \times \ldots P_{i_n,j_n} \times A \times P_{i_n,j_n} \ldots \times P_{i_1,j_1} \times P_{i_1,j_1} \times P_{i_2,j_2} \times \ldots P_{i_n,j_n} \times B \times P_{i_n,j_n} \ldots \times P_{i_1,j_1} \\ &= P_{i_1,j_1} \times P_{i_2,j_2} \times \ldots P_{i_n,j_n} \times A \times B \times P_{i_n,j_n} \ldots \times P_{i_1,j_1} \\ &= P_{i_1,j_1} \times P_{i_2,j_2} \times \ldots P_{i_n,j_n} \times I \times P_{i_n,j_n} \ldots \times P_{i_1,j_1} \\ &= I \end{aligned}$$

The next lemma is ${A'_1}^{-1} = B'_1 \ - \ B'_2{B'_3}^{-1}B'_4$. Consider a matrix X. X is partitioned into 4 parts as listed below.

$$X = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

Thus X's inverse will be

$$\begin{aligned} X^{-1} &= \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}^{-1} = \begin{bmatrix} {A_1}^{-1} + {A_1}^{-1}A_2(A_4 - A_3{A_1}^{-1}A_2)^{-1}A_3{A_1}^{-1} & -{A_1}^{-1}A_2(A_4 - A_3{A_1}^{-1}A_2)^{-1} \\ -(A_4 - A_3{A_1}^{-1}A_2)^{-1}A_3{A_1}^{-1} & (A_4 - A_3{A_1}^{-1}A_2)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} \end{aligned}$$

So that ${A'_1}^{-1} = B'_1 \ - \ B'_2{B'_3}^{-1}B'_4$.