Covariance matrix for coefficients is positive definition

Uma Roy

October 11, 2018

Theorem 1. Let Σ be the covariance matrix of a Gaussian graphical model on p nodes (i.e. $\Sigma \in \mathbb{R}^{p\times p}$) and let $N_1, N_2, \ldots, N_S$ be subsets of $[p]$ such that for all subsets N_t, it is of the form $N_t = \{k, l, \text{neigh}(k), r_t\}$, where r_t is a random vertex $r_t \neq k, l, \text{neigh}(k)$ that satisfies the following property: *There exists at least 1 path from l to r not going through k or neigh(k).*

Given a matrix M, we let M_{N_t,N_q} denote the submatrix obtained by taking all rows corresponding to nodes in N_t and all columns in N_q. We also let S denote the empirical covariance matrix of samples drawn from our Gaussian graphical model.

Let $\vec{X}$ be the vector where the t-th component $X_t = (S_{N_t,N_t})^{-1}_{k,l}$, where we abuse notation a bit by letting the entry (k,l) of this matrix correspond to the position of (k,l) in N_t (as opposed to the position of (k,l) in $[p]$, as would be normal). We see that $X_t = -\rho^{k,l}_{\{\text{neigh}(k),r_t\}} \cdot \sqrt{\sigma^{kk}\sigma^{ll}}$ (where $\sigma^{kk} = (\Sigma)^{-1}_{kk}$) i.e. the negative partial covariance between nodes k, l when conditioning on neigh$(k), r_t$. Cov$(\vec{X})$ is an important quantity because it reveals how the coefficient of k is correlated across the different regressions of node k on subsets N_t.

This theorem shows that Cov$(\vec{X})$ is positive definite (full-rank).

Proof. By Taylor expansion, we have

$$(S_{N_t,N_t})^{-1} = (\Sigma_{N_tN_t})^{-1} + \text{ISS}\left((\Sigma_{N_tN_t})^{-1}\right)(S_{N_tN_t} - \Sigma_{N_tN_t}).$$

Let $S_{N_tN_t} = D_t \cdot S$. We note that we are treating $S_{N_tN_t}$ as a vector $\in \mathbb{R}^d$ and $S \in \mathbb{R}^{d'}$ as a vector, where $d = \frac{|N_t|(|N_t|+1)}{2}$ and $d' = \frac{p(p+1)}{2}$. Then $D_t \in R^{d\times d'}$.

We also know that

$$\lim_{n\to\infty}(S - \Sigma) \to \mathcal{N}(0, G),$$

where $G \in \mathbb{R}^{d'\times d'}$ and $G = \text{ISS}(\Sigma)$.

We note that

$$\mathbb{E}\left[(S_{N_t,N_t})^{-1}\right] = (\Sigma_{N_tN_t})^{-1},$$

since $\mathbb{E}[S - \Sigma] = \vec{0}$.

Then to compute Cov(X_t, X_q), we must compute the following:

$$\mathbb{E}\left[\left((S_{N_t,N_t})^{-1} - (\Sigma_{N_t,N_t})^{-1}\right)\left((S_{N_q,N_q})^{-1} - (\Sigma_{N_q,N_q})^{-1}\right)^T\right]$$

$$= \mathbb{E}\left[\left(\text{ISS}\left((\Sigma_{N_tN_t})^{-1}\right)(S_{N_tN_t} - \Sigma_{N_tN_t})\right)\left(\text{ISS}\left((\Sigma_{N_qN_q})^{-1}\right)(S_{N_qN_q} - \Sigma_{N_qN_q})\right)^T\right]$$

Let $I_t = \text{ISS}\left((\Sigma_{N_tN_t})^{-1}\right)$ and I_q similarly.

$$= \mathbb{E}\left[(I_t \cdot D_t(S - \Sigma))(I_q \cdot D_q(S - \Sigma))^T\right]$$

$$= \mathbb{E}\left[I_t \cdot D_t(S - \Sigma)(S - \Sigma)^T D_q^T I_q^T\right]$$

$$= I_t \cdot D_t \cdot G \cdot D_q^T \cdot I_q^T.$$

Thus we can compute the (t,q)-th entry of $\mathrm{Cov}(\vec{X})$ as follows:

$$\mathrm{Cov}(\vec{X})_{t,q} = \left(I_t \cdot D_t \cdot G \cdot D_q^T \cdot I_q^T\right)_{(k,l),(k,l)},$$

where we again abuse notation and take the entry of the resulting matrix that corresponds position of entry (k,l) in (S_{N_t,N_t}) (when we treat it as a vector $\in \mathbb{R}^d$) and similarly for the position of (k,l) in $\left(S_{N_q,N_q}\right)$.

Because of the matrix identity: $\left(ABC^T\right)_{i,j} = A_i B (C_j)^T$, where A_i is the i-th row of A, if we let $I'_t := I_t \cdot D_t$, then we can rewrite

$$\mathrm{Cov}(\vec{X})_{t,q} = \left(I_t \cdot D_t \cdot G \cdot D_q^T \cdot I_q^T\right)_{(k,l),(k,l)} = (I'_t)_{(k,l)} G \left((I'_q)_{(k,l)}\right)^T,$$

where by yet another abuse of notation, $(I'_t)_{(k,l)}$ denotes the row of I'_t that corresponds to the pair (k,l) in S_{N_t,N_t} (regarded as a vector $\in \mathbb{R}^d$).

If we let $\mathcal{I} \in \mathbb{R}^{S,d'}$ (where S is the number of susbets we're considering) be the matrix where each row $\mathcal{I}_t = (I'_t)_{(k,l)}$, again using the same matrix identity as before, we have

$$\mathrm{Cov}(\vec{X}) = \mathcal{I} G \mathcal{I}^T.$$

We know that G is full-rank because it is an Isserlis matrix. Thus to show $\mathrm{Cov}(\vec{X})$ is positive definite, it suffices to show that $v^T \mathcal{I} \neq \vec{0}$ for $v \neq \vec{0}$, i.e. $\mathcal{I}$ has full row-rank.

To show that $\mathcal{I}$ has full row-rank, we show that the rows of $\mathcal{I}$ are all linearly independent. In particular, we show that for each row of $\mathcal{I}$, there is at least 1 entry that is non-zero *only* in that row, which implies the rows are linearly independent.

Recall that the t-th row of $\mathcal{I}$ is: $(I_t \cdot D_t)_{(k,l)} = (I_t)_{(k,l)} \cdot D_t$. Effectively D_t is transforming entries of I_t to the corresponding indexing used for all nodes (since I_t is indexed by only the nodes in N_t). In particular, if we let $\mathcal{I}_{t,(u,v)}$ denote the entry in row t of $\mathcal{I}$ that corresponds to the pair of nodes (u,v), then

$$\mathcal{I}_{t,(u,v)} = \begin{cases} (I_t)_{(k,l),(u,v)}, & \text{if } (u,v) \in N_t \times N_t \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

In particular for each t, consider the entry $\mathcal{I}_{t,(k,r_t)}$. By the formula for Isserlis matrices, we have that $\mathcal{I}_{t,(k,r_t)} = (I_t)_{(k,l),(k,r_t)} = \sigma^{kk}_{N_t}\sigma^{lr_t}_{N_t} + \sigma^{kr_t}_{N_t}\sigma^{lk}_{N_t} = \sigma^{kk}_{N_t}\sigma^{lr_t}_{N_t}$, since $\sigma^{lk}_{N_t} = 0$, since all paths from k to l must go through $\mathrm{neigh}(k) \in N_t$. However, since we imposed the condition that there is a path from l to r_t that does not go through k or $\mathrm{neigh}(k)$, we have that $\sigma^{lr_t}_{N_t} \neq 0$ and since $\sigma^{kk}_{N_t} \neq 0$, we have that $\mathcal{I}_{t,(k,r_t)} \neq 0$. But consider the same entry in any other row, i.e. $\mathcal{I}_{q,(k,r_t)}$, for $q \neq t$. Then we have that since $r_t \notin N_q$, this entry is necessarily $= 0$. Thus we have shown each row of $\mathcal{I}$ has at least 1 entry that is non-zero in *only* that row, proving that it has full row-rank, which completes the proof.

□