An Innovative Approach for Analysing Rank Deficient Covariance Matrices

Gabriel H. Tucci

In collaboration with Ke Wang

ISIT 2012 @ MIT

July 5 2012
Motivation: Covariance Estimates

We have $m$ random variables

Correlation between random variables? from $n$ observations

- Weather Forecast: Sensor Network where sensors are measuring temperature, pressure, etc
- Military Applications (adaptative sensor array)
- Gene Expression Arrays
- High Dimensional Problems with $n \leq m$

We can’t perform as many observations as the number of variables!
Mathematical Formulation

Let \( \{x_1, x_2, \ldots, x_n\} \) be \( n \) observations of a \( m \)-dimensional random vector \( X \in \mathbb{R}^{m \times 1} \). For simplicity, suppose \( X \) has mean zero.

True covariance Matrix: \( \Sigma = \text{Cov}(X) = \mathbb{E}(XX^T) \)

Sample covariance Matrix: \( K = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \) converges to \( \Sigma \) as \( n \to \infty \)
Let \( \{x_1, x_2, \ldots, x_n\} \) be \( n \) observations of a \( m \)-dimensional random vector \( X \in \mathbb{R}^{m \times 1} \). For simplicity, suppose \( X \) has mean zero.

True covariance Matrix: \[ \Sigma = \text{Cov}(X) = \mathbb{E}(XX^T) \]

Sample covariance Matrix: \[ K = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \] converges to \( \Sigma \) as \( n \to \infty \)

Recover \( \Sigma \) or \( \Sigma^{-1} \) from the sample covariance matrix \( K \) when the information is less than the dimension:\[ n \leq m \]

- \( K \) is singular and has at least \( m - n \) zero eigenvalues
Let \( \{x_1, x_2, \ldots, x_n\} \) be \( n \) observations of a \( m \)-dimensional random vector \( X \in \mathbb{R}^{m \times 1} \). For simplicity, suppose \( X \) has mean zero.

True covariance Matrix: \( \Sigma = \text{Cov}(X) = \mathbb{E}(XX^T) \)

Sample covariance Matrix: \( K = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^\ast \) converges to \( \Sigma \) as \( n \to \infty \)

Recover \( \Sigma \) or \( \Sigma^{-1} \) from the sample covariance matrix \( K \) when the information is less than the dimension:
\[
 n \leq m
\]

- \( K \) is singular and has at least \( m - n \) zero eigenvalues
- In many applications, we need \( \Sigma^{-1} \) (e.g. linear prediction)
Mathematical Formulation

Let \( \{x_1, x_2, \ldots, x_n\} \) be \( n \) observations of a \( m \)--dimensional random vector \( X \in \mathbb{R}^{m \times 1} \). For simplicity, suppose \( X \) has mean zero.

True covariance Matrix: \( \Sigma = \text{Cov}(X) = \mathbb{E}(XX^T) \)

Sample covariance Matrix: \( \hat{K} = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^* \) converges to \( \Sigma \) as \( n \to \infty \)

Recover \( \Sigma \) or \( \Sigma^{-1} \) from the sample covariance matrix \( \hat{K} \) when the information is less than the dimension:

\( n \leq m \)

- \( \hat{K} \) is singular and has at least \( m - n \) zero eigenvalues
- In many applications, we need \( \Sigma^{-1} \) (e.g. linear prediction)

Classical solution: ridge regression method or diagonal loading

\[ \alpha \hat{K} + \beta I_m \]

where \( \alpha, \beta > 0 \). Ledoit and Wolf’s result to choose the optimal parameters.
Method 1: Invcov\(_{p}\) Estimation

In a joint paper [Marzetta, T., Simon] we suggested a new approach to estimate \(\Sigma\) or \(\Sigma^{-1}\) from the sample covariance matrix \(K\)

Fix a parameter \(p \leq n\) (to be tuned later) and consider the Stiefel manifold

\[
\Omega_{p,m} := \left\{ \Phi \in M_{p,m}(\mathbb{C}) : \Phi\Phi^* = I_p \right\}
\]

with isotropically random probability measure
Method 1: Invcov\(p\) Estimation

In a joint paper [Marzetta, T., Simon] we suggested a new approach to estimate \(\Sigma\) or \(\Sigma^{-1}\) from the sample covariance matrix \(K\).

Fix a parameter \(p \leq n\) (to be tuned later) and consider the Stiefel manifold

\[
\Omega_{p,m} := \left\{ \Phi \in M_{p,m}(\mathbb{C}) : \Phi\Phi^* = I_p \right\}
\]

with isotropically random probability measure

Haar Integration Method

\[
\text{cov}_p(K) := E\left( \Phi^* (\Phi K \Phi^*) \Phi \right) \approx \Sigma
\]

\[
\text{invcov}_p(K) := E\left( \Phi^* (\Phi K \Phi^*)^{-1} \Phi \right) \approx \Sigma^{-1}
\]
What is already known about $\text{cov}_p$ and $\text{invcov}_p$

- $\text{cov}_p(K) = \frac{1}{m^2-1} \left( (mp - 1) \cdot K + (m - p) \cdot \text{tr}(K) \cdot I_m \right)$ which is equivalent to diagonal loading!

- If $K = U D U^*$, where $U$ is unitary and $D = \text{diag}(d_1, \ldots, d_n, 0_{m-n})$, then $\text{invcov}_p(K) = U \text{invcov}_p(D) U^*$, and $\text{invcov}_p(D) = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n, \mu I_{m-n})$.

- Closed-form expression for $\lambda_k$ and $\mu$

- Functional expression relating the parameters

- Asymptotic expressions using Free Probability Theory
What is already known about $\text{cov}_p$ and $\text{invcov}_p$

- $\text{cov}_p(K) = \frac{1}{m^2-1} \left( (mp - 1) \cdot K + (m - p) \cdot \text{tr}(K) \cdot I_m \right)$ which is equivalent to diagonal loading!

- If $K = U D U^*$, where $U$ is unitary and $D = \text{diag}(d_1, \ldots, d_n, 0_{m-n})$, then
  
  $$\text{invcov}_p(K) = U \text{invcov}_p(D) U^*,$$

  and
  
  $$\text{invcov}_p(D) = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n, \mu I_{m-n}).$$
What is already known about $\text{cov}_p$ and $\text{invcov}_p$

- $\text{cov}_p(K) = \frac{1}{m^2-1} \left( (mp - 1) \cdot K + (m - p) \cdot \text{tr}(K) \cdot I_m \right)$ which is equivalent to diagonal loading!

- If $K = U D U^*$, where $U$ is unitary and $D = \text{diag}(d_1, \ldots, d_n, 0_{m-n})$, then

  $$\text{invcov}_p(K) = U \text{invcov}_p(D) U^*,$$

  and

  $$\text{invcov}_p(D) = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n, \mu I_{m-n}).$$

- Closed-form expression for $\lambda_k$ and $\mu$

- Functional expression relating the parameters

- Asymptotic expressions using Free Probability Theory
Other Mathematical Properties

We continue to explore the mathematical properties related to \( \text{invcov}_p \)

**Theorem (T., Wang)**

For \( m \times m \) Hermitian matrix \( K \)

\[
\text{invcov}_p(K) \in \mathcal{A}(K) = \left\{ \alpha_{m-1}K^{m-1} + \alpha_{m-2}K^{m-2} + \ldots + \alpha_1K + \alpha_0I_m : \alpha_i \in \mathbb{C} \right\}
\]
Other Mathematical Properties

We continue to explore the mathematical properties related to $\text{invcov}_p$

**Theorem (T., Wang)**

For $m \times m$ Hermitian matrix $K$

$$\text{invcov}_p(K) \in \mathcal{A}(K) = \left\{ \alpha_{m-1}K^{m-1} + \alpha_{m-2}K^{m-2} + \ldots + \alpha_1K + \alpha_0I_m : \alpha_i \in \mathbb{C} \right\}$$

**Theorem (T., Wang)**

For every positive integer $N$

$$\mathbb{E}\left(\Phi^*(\Phi K \Phi^*)^N \Phi\right) = \sum_{i=0}^{N} a_i K^i$$

We use results from representation theory to provide explicit formulas to compute the coefficients $a_k$'s.

For example,

$$\mathbb{E}(\Phi(\Phi^* K \Phi)^2 \Phi^*) = (c_0 + c_1 + c_2)K^2 + (c_0 - c_2)\text{tr}(K)K$$

$$+ \left( c_0 \frac{\text{tr}(K)^2 + \text{tr}(K^2)}{2} - c_1 + c_2 \frac{\text{tr}(K)^2 - \text{tr}(K^2)}{2} \right) I_n$$

where $c_0 = \frac{1}{4} \frac{(3+p)!(m-1)!}{(3+m)!(p-1)!}$, $c_1 = \frac{1}{4} \frac{(2+p)!(m-2)!}{(2+m)!(p-2)!}$, $c_2 = \frac{1}{4} \frac{(1+p)!(m-3)!}{(1+m)!(p-3)!}$. 

A permutation is a bijection $\sigma : \{1, 2, \ldots, m\} \rightarrow \{1, 2, \ldots, m\}$. We can visualize $\sigma$ with a directed graph $G$ on $m$ vertices.
A permutation is a bijection $\sigma : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$. We can visualize $\sigma$ with a directed graph $G$ on $m$ vertices.

For example, $m = 6$

\[
\begin{array}{ccccccc}
1 & \rightarrow & 2 & \rightarrow & 4 & \rightarrow & 3 \\
\rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \\
3 & \rightarrow & 5 & \rightarrow & 6 & \rightarrow &
\end{array}
\]

$K(\sigma) =$ number of connected components in $G_\sigma$. 

\begin{align*}
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{align*}
Method 2: Random permutations under Ewens measure

A permutation is a bijection $\sigma: \{1, 2, \ldots, m\} \rightarrow \{1, 2, \ldots, m\}$. We can visualize $\sigma$ with a directed graph $G$ on $m$ vertices.

For example, $m = 6$

We associate $\sigma$ with a permutation matrix $M_{\sigma}$, which is an $m \times m$ unitary matrix such that

$$M_{\sigma}(i, j) = 1_{\sigma(i) = j}$$

Hence

$$M_{\sigma} = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$
Let $S_m$ be the set of all permutations on $\{1, 2, \ldots, m\}$. We endow $S_m$ with \textbf{Ewens measure} by choosing $\sigma$ with probability

$$p_{\theta,m}(\sigma) = \frac{\theta^{K(\sigma)}}{\theta(\theta + 1)\ldots(\theta + m - 1)}$$

where $\theta > 0$ and $K(\sigma)$ is the number of connected components in the graph $G_{\sigma}$. 
Method 2: Random permutation under Ewens measure

Let $S_m$ be the set of all permutations on $\{1, 2, \ldots, m\}$

We endow $S_m$ with Ewens measure by choosing $\sigma$ with probability

$$p_{\theta,m}(\sigma) = \frac{\theta^K(\sigma)}{\theta(\theta + 1)\ldots(\theta + m - 1)}$$

where $\theta > 0$ and $K(\sigma)$ is the number of connected components in the graph $G_\sigma$

Specially, for $\theta = 1$, it is the uniform measure on $S_m$, which gives the same weight to every permutation $\sigma$
Method 2: Ewens Estimation

Recall $\Sigma$ is the true covariance matrix and $K$ is the sample covariance matrix generated from $n$ observations.

Define

$$K_\theta = \mathbb{E}(M_\sigma K M_\sigma^*) \approx \Sigma$$
Method 2: Ewens Estimation

Recall $\Sigma$ is the true covariance matrix and $K$ is the sample covariance matrix generated from $n$ observations.

Define

$$K_\theta = \mathbb{E}(M_\sigma K M^*_\sigma) \approx \Sigma$$

Theorem (T.,Wang)

*We provide a closed-form expression for*

$$K_\theta = \mathbb{E}(M_\sigma K M^*_\sigma).$$
Method 2: Ewens Estimation

Recall $\Sigma$ is the true covariance matrix and $K$ is the sample covariance matrix generated from $n$ observations.

Define

$$K_\theta = \mathbb{E}(M_\sigma KM_\sigma^*) \approx \Sigma$$

**Theorem (T.,Wang)**

*We provide a closed-form expression for*

$$K_\theta = \mathbb{E}(M_\sigma KM_\sigma^*).$$

If $\theta = 1$,

$$K_1 = \frac{\alpha ee^T}{m} + \beta(I_m - \frac{ee^T}{m})$$

where $\alpha = \frac{\sum_{i,j=1}^m K_{ij}}{m}$ and $\beta = \frac{\text{tr}(K) - \alpha}{m - 1}$. 


Method 2: Ewens Estimation

Recall $\Sigma$ is the true covariance matrix and $K$ is the sample covariance matrix generated from $n$ observations.

Define

$$K_\theta = \mathbb{E}(M\sigma KM_\theta^*) \approx \Sigma$$

**Theorem (T., Wang)**

We provide a closed-form expression for

$$K_\theta = \mathbb{E}(M\sigma KM_\theta^*).$$

If $\theta = 1$,

$$K_1 = \alpha \frac{ee^T}{m} + \beta (I_m - \frac{ee^T}{m})$$

where $\alpha = \frac{\sum_{i,j=1}^{m} K_{ij}}{m}$ and $\beta = \frac{\text{tr}(K) - \alpha}{m - 1}$.

If $K = D = \text{diag}(d_1, \ldots, d_m)$ then

$$K_\theta = \frac{\theta - 1}{\theta + m - 1} D + \frac{\text{tr}(D)}{\theta + m - 1} I_m,$$

which is equivalent to diagonal loading.
Concrete Examples: Power Toeplitz matrix

Consider $m \times m$ power Toeplitz matrix

$$A_\alpha = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{m-1} \\ \alpha & 1 & \alpha & \cdots & \alpha^{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{m-2} & \alpha^{m-3} & \cdots & 1 & \alpha \\ \alpha^{m-1} & \alpha^{m-2} & \cdots & \alpha & 1 \end{pmatrix} = (\alpha|i-j|)_{1 \leq i,j \leq m}.$$
Concrete Examples: Power Toeplitz matrix

Consider $m \times m$ power Toeplitz matrix

$$A_{\alpha} = \begin{pmatrix}
1 & \alpha & \alpha^2 & \cdots & \alpha^{m-1} \\
\alpha & 1 & \alpha & \cdots & \alpha^{m-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\alpha^{m-2} & \alpha^{m-3} & \cdots & 1 & \alpha \\
\alpha^{m-1} & \alpha^{m-2} & \cdots & \alpha & 1
\end{pmatrix} = (\alpha^{i-j})_{1 \leq i,j \leq m}.$$  

Theorem

- $\det(A_{\alpha}) = (1 - \alpha^2)^{m-1}$.
- $A_{\alpha} \geq 0$ if and only if $|\alpha| \leq 1$.
- For $\alpha \neq 1$,

$$A_{\alpha}^{-1} = \frac{1}{1 - \alpha^2} \begin{pmatrix}
1 & -\alpha & -\alpha & \cdots & -\alpha \\
-\alpha & 1 + \alpha^2 & -\alpha & \cdots & -\alpha \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
-\alpha & -\alpha & 1 + \alpha^2 & -\alpha & \cdots \\
-\alpha & -\alpha & -\alpha & \cdots & 1
\end{pmatrix}.$$
For an $m \times m$ matrix $A$, the normalized Frobenius norm $\|A\|_2 = \frac{\sqrt{\text{Tr}(AA^*)}}{m}$. The true covariance matrix is $\Sigma = A_\alpha$ and the sample covariance matrix is $K$, recall
Method 1 for Power Toeplitz Matrices

For an $m \times m$ matrix $A$, the normalized Frobenius norm $\|A\|_2 = \sqrt{\text{Tr}(AA^*)}/m$. The true covariance matrix is $\Sigma = A_\alpha$ and the sample covariance matrix is $K$, recall We compute the Frobenius norms

$$f(m, n, \alpha, p) = \|\Sigma - \text{invcov}_p(K)^{-1}\|_2 \quad g(m, n, \alpha, p) = \|\Sigma^{-1} - \text{invcov}_p(K)\|_2$$

and try to explore the optimal values of $p$ that minimize these norms
For an $m \times m$ matrix $A$, the normalized Frobenius norm $\|A\|_2 = \sqrt{\text{Tr}(AA^*)}/m$. The true covariance matrix is $\Sigma = A_\alpha$ and the sample covariance matrix is $K$, recall We compute the Frobenius norms

$$f(m, n, \alpha, p) = \|\Sigma - \text{invcov}_p(K)^{-1}\|_2 \quad g(m, n, \alpha, p) = \|\Sigma^{-1} - \text{invcov}_p(K)\|_2$$

and try to explore the optimal values of $p$ that minimize these norms.
Simulation: Method 1 for Power Toeplitz Matrices

\[ f(m, n, \alpha, p) = \| A_\alpha - \text{invcov}_p(K)^{-1} \|_2 \]
Simulation: Method 2 for Power Toeplitz Matrices

$$F(m, n, \alpha, \theta) = \|A_\alpha - \mathbb{E}(M_\sigma KM_\sigma^*)\|_2$$

Ewens Estimation of the Covariance Matrix $m = 200$ and $\alpha = 0.5$

- $n = 40$, optimal $\theta = 90$ with MSE = 0.7635
- $n = 80$, optimal $\theta = 174$ with MSE = 0.7179
- $n = 150$, optimal $\theta = 270$ with MSE = 0.6607
- $n = 180$, optimal $\theta = 337$ with MSE = 0.6405
Comparison Between the Different Methods

- True Covariance Matrix with $m=200$ and $\alpha = 0.5$
- Sample Covariance with $n = 150$
- Inverse of invcov with optimum $p$ ($p = 45$) MSE = 0.7420
- Ewens Method with Optimum $\theta$ ($\theta = 261$) MSE = 0.6607
Comparison Between the Different Methods

Forbenius Norm of invcov and inverse of invcov

Parameter p

- Inverse invcov : optimum p = 45 MSE = 0.7420
- invcov : optimum p = 40 MSE = 0.8942

Ewens Estimation of the Covariance Matrix m = 200, n = 150 and $\alpha = 0.5$

Optimum $\theta = 270$ with MSE = 0.8607
Method 3

Let $p \leq m$ and consider a set $S_{p,m} = \{\sigma : \{1, \ldots, p\} \to \{1, \ldots, m\} \mid \sigma \text{ injection}\}$. If $p = m$ then $S_{p,m}$ is the symmetric group $S_m$. For every $\sigma \in S_{p,m}$ define a $p \times m$ matrix

$$U_\sigma = \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(p)} \end{pmatrix}.$$  

Let $\Omega_\sigma = \{\tilde{\sigma} \in S_m : \tilde{\sigma}_{\{1,\ldots,p\}} = \sigma\}$. Recall $p_{\theta,m}$ is the Ewens measure on $S_m$. 

Let $p \leq m$ and consider a set $S_{p,m} = \{\sigma : \{1, \ldots, p\} \rightarrow \{1, \ldots, m\} \mid \sigma \text{ injection}\}$. If $p = m$ then $S_{p,m}$ is the symmetric group $S_m$. For every $\sigma \in S_{p,m}$ define a $p \times m$ matrix

$$U_\sigma = \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(p)} \end{pmatrix}.$$ 

Let $\Omega_\sigma = \{\tilde{\sigma} \in S_m : \tilde{\sigma}\{1,\ldots,p\} = \sigma\}$. Recall $p_{\theta,m}$ is the Ewens measure on $S_m$. We sample $\sigma \in S_{p,m}$ randomly with probability

$$\mu_{\theta,m,p}(\sigma) = p_{\theta,m}(\Omega_\sigma) = \sum_{\tilde{\sigma} \in \Omega_\sigma} p_{\theta,m}(\sigma).$$

$$K_\theta = \mathbb{E}\left(U_\sigma^*(U_\sigma KU_\sigma^*)U_\sigma\right) \approx \Sigma$$

$$\tilde{K}_\theta = \mathbb{E}\left(U_\sigma^*(U_\sigma KU_\sigma^*)^+U_\sigma\right) \approx \Sigma^{-1}$$

where $(U_\sigma KU_\sigma^*)^+$ is the Moore–Penrose pseudoinverse of matrix $U_\sigma KU_\sigma^*$. 

Conclusion

- Mathematical properties related to $\text{invcov}_p$
Conclusion

- Mathematical properties related to $\text{invcov}_p$
- Explicit formula for $\mathbb{E}(M_\sigma K M_\sigma^*)$
Conclusion

- Mathematical properties related to $\text{inv}cov_P$
- Explicit formula for $E(M_\sigma K M^*_\sigma)$
- Simulations to investigate and compare both methods
Conclusion

- Mathematical properties related to $\text{invcov}_p$
- Explicit formula for $\mathbb{E}(M_\sigma K M_\sigma^*)$
- Simulations to investigate and compare both methods
- These methods are easy to implement + do not assume previous knowledge on the covariance
Conclusion

- Mathematical properties related to $\text{inv} \text{cov}_p$

- Explicit formula for $\mathbb{E}(M_\sigma K M_\sigma^*)$

- Simulations to investigate and compare both methods

- These methods are easy to implement + do not assume previous knowledge on the covariance

- A new method that combines the ideas of the previous two methods
Conclusion

- Mathematical properties related to $\text{invcov}_p$
- Explicit formula for $\mathbb{E}(M_\sigma KM_\sigma^*)$
- Simulations to investigate and compare both methods
- These methods are easy to implement + do not assume previous knowledge on the covariance
- A new method that combines the ideas of the previous two methods

Thanks!
Appendix: Explicit formula for $\mathbb{E}_\theta(M\sigma AM^*_\sigma)$

Theorem

Let $A = (a_{ij})$ be an $N \times N$ normal matrix. Then $\mathbb{E}_\theta(M\sigma AM^*_\sigma) = B_\theta = (b_{ij})_{1 \leq i, j \leq N}$, where

$$b_{ii} = \frac{\theta - 1}{\theta + N - 1} a_{ii} + \frac{1}{\theta + N - 1} \text{tr}(A),$$

for $i \neq j$, $b_{ij} = \frac{1}{(\theta + N - 2)(\theta + N - 1)} \left( \sum_{l \neq k} a_{lk} + (\theta - 1) \sum_{k \neq i, j} (a_{ik} + a_{kj}) + (\theta^2 - 1)a_{ij} \right)$.
Appendix: Explicit formula for $\mathbb{E}_\theta(M\sigma AM_\sigma^*)$

**Theorem**

Let $A = (a_{ij})$ be an $N \times N$ normal matrix. Then $\mathbb{E}_\theta(M\sigma AM_\sigma^*) = B_\theta = (b_{ij})_{1 \leq i,j \leq N}$, where

$$b_{ii} = \frac{\theta - 1}{\theta + N - 1} a_{ii} + \frac{1}{\theta + N - 1} \text{tr}(A),$$

(3.1)

for $i \neq j$, $b_{ij} = \frac{1}{(\theta + N - 2)(\theta + N - 1)} \left( \sum_{l \neq k} a_{lk} + (\theta - 1) \sum_{k \neq i,j} (a_{ik} + a_{kj}) + (\theta^2 - 1)a_{ij} \right)$. (3.2)

$$B_\theta = \frac{(\theta - 1)^2}{(\theta + N - 2)(\theta + N - 1)} A + \frac{(\theta - 1)(N - 1)}{(\theta + N - 2)(\theta + N - 1)} \text{diag}(a_{11}, a_{22}, \ldots, a_{NN})$$

$$+ \frac{\text{tr}(A)}{\theta + N - 2} I_N + \frac{eAe^T}{(\theta + N - 2)(\theta + N - 1)} ee^T + \frac{\theta - 1}{(\theta + N - 2)(\theta + N - 1)} K_N,$$

(3.3)

where $K_N$ has diagonal entries 0 and $(K_N)_{ij} = \sum_{k \neq i} a_{ik} + \sum_{k \neq j} a_{jk}$. 