Estimation of Low Rank and Sparse Covariance Matrices

Szymon Czop

Wrocław University
Data Science

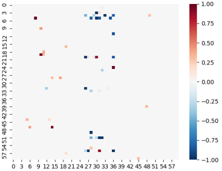
29 września 2021

Agenda

What connections do we look for?

We are looking for matrices that are **sparse** and has **low rank**.

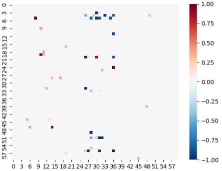
- Sparse, full rank
- Low rank not sparse

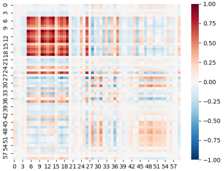


What connections do we look for?

We are looking for matrices that are **sparse** and has **low rank**.

- Sparse, full rank
- Low rank not sparse
- Low rank and sparse !

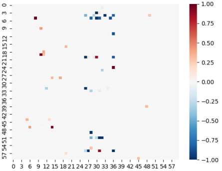


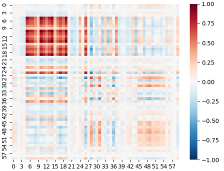


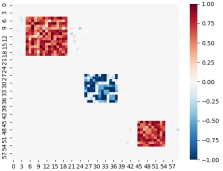
What connections do we look for?

We are looking for matrices that are **sparse** and has **low rank**.

- Sparse, full rank
- Low rank not sparse
- Low rank and sparse !







Introducing **Low-Rank Sparse Estimator** (LoRSEr)

$$\hat{B} := \underset{B}{\operatorname{argmin}} \left\{ \frac{1}{2}||Y - B||_F^2 + \lambda_N \|B\|_* + \lambda_L \|\operatorname{vec}(W \circ B)\|_1 \right\}$$

Where:

- Y - empirical covariance matrix.
- $||Y - B||_F$ - Frobenius norm.
- $||B||_*$ - nuclear norm responsible for low rank.
- W - matrix of weights.
- $\|\operatorname{vec}(W \circ B)\|_1$ - l_1 norm responsible for sparsity.
- λ_N, λ_L - scales fine-tuned in the CV process.

Alternating Direction Method of Multipliers (ADMM)

Numerical Solution

$$F(B) := \underbrace{\frac{1}{2}||Y - B||_F^2}_{f(B)} + \underbrace{\lambda_N \|B\|_*}_{g(B)} + \underbrace{\lambda_L \|\operatorname{vec}(W \circ B)\|_1}_{h(B)}.$$

Alternating Direction Method of Multipliers (ADMM)

Low - **R**ank **S**parse **E**stimator (LoRSEr)

$$B^{[k+1]} := \underset{B}{\operatorname{argmin}} \left\{ 2f(B) \ + \ \delta_1^{[k]} \left\| D^{[k]} + \frac{Z_1^{[k]}}{\delta_1^{[k]}} - B \right\|_F^2 \right\}$$

$$C^{[k+1]} := \underset{C}{\operatorname{argmin}} \left\{ 2g(C) \ + \ \delta_2^{[k]} \left\| D^{[k]} + \frac{Z_2^{[k]}}{\delta_2^{[k]}} - C \right\|_F^2 \right\}$$

$$D^{[k+1]} := \underset{D}{\operatorname{argmin}} \left\{ 2h(D) \ + \ \delta_1^{[k]} \left\| D + \frac{Z_1^{[k]}}{\delta_1^{[k]}} - B^{[k+1]} \right\|_F^2 + \delta_2^{[k]} \left\| D + \frac{Z_2^{[k]}}{\delta_2^{[k]}} - C^{[k+1]} \right\|_F^2 \right\}$$

$$\left\{ \begin{array}{l} Z_1^{[k+1]} := \ Z_1^{[k]} + \delta_1^{[k]} \left(D^{[k+1]} - B^{[k+1]} \right) \\ Z_2^{[k+1]} := \ Z_2^{[k]} + \delta_2^{[k]} \left(D^{[k+1]} - C^{[k+1]} \right) \end{array} \right. .$$

Properties of LoRSEr

Properties

Property 1

For any pair of the tuning parameters, there **exists** solution to optimisation problem and this solution is **unique**.

Properties of LoRSEr

Property 1

For any pair of the tuning parameters, there **exists** solution to optimisation problem and this solution is **unique**.

Property 2

Solution is symmetric

Properties of LoRSEr

Property 1

For any pair of the tuning parameters, there **exists** solution to optimisation problem and this solution is **unique**.

Property 2

Solution is symmetric.

Property 3

Solution is invariable under reordering applied to rows and columns of Y matrix.

Experiments (Generating data and loss calculation)

X - experiment matrix
X_i - i-th row of matrix

$$X_i \sim \mathcal{N}(\mu, \Sigma) .$$

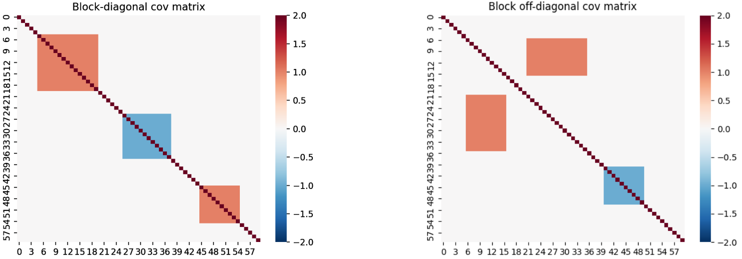
Where μ is a zero vector and Σ is predefined covariance matrix we want to estimate after data generation.

Loss Function

$$\text{MSEr} = ||\hat{B} - \Sigma||_{F^\star}^2 / ||\Sigma||_{F^\star}^2$$

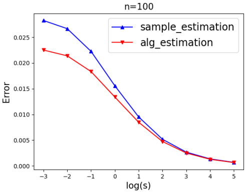
where $||.||_*$ is Frobenius Norm of matrix with diagonal entries excluded.

Experiments (Examples of covariance matrices)

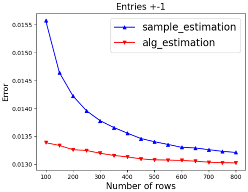


Rysunek 1: Examples of covariance matrices we want to recover from data

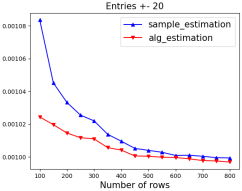
Differences in performance



(a) Different values

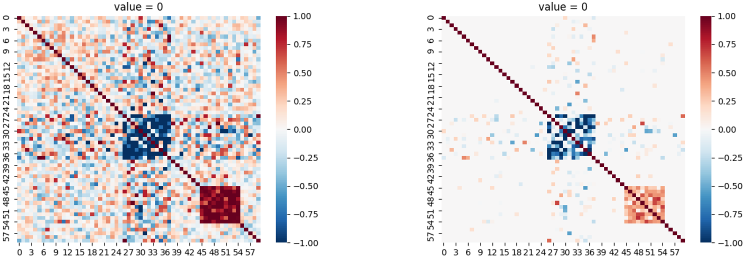


(b) Small values



(c) Big values

Differences in outcome



Rysunek 3: Examples of covariance matrices recovered from data.

End

Thank You for Your Attention

Alternating Direction Method of Multipliers (ADMM)

$$F(B) := \underbrace{\frac{1}{2}||Y - B||_F^2}_{f(B)} + \underbrace{\lambda_N \|B\|_*}_{g(B)} + \underbrace{\lambda_L \|\operatorname{vec}(W \circ B)\|_1}_{h(B)}.$$

Introducing variables C and D

$$\underset{B,C,D}{\operatorname{argmin}} \; \{f(B) + g(C) + h(D)\} \qquad \text{such that} \; \begin{cases} D - B = 0 \\ D - C = 0 \end{cases}$$

Alternating Direction Method of Multipliers (ADMM)

$$F(B) := \underbrace{\frac{1}{2}||Y - B||_F^2}_{f(B)} + \underbrace{\lambda_N \|B\|_*}_{g(B)} + \underbrace{\lambda_L \| \mathrm{vec}(W \circ B)\|_1}_{h(B)}.$$

Introducing variables C and D

$$\underset{B,C,D}{\operatorname{argmin}} \ \{f(B) + g(C) + h(D)\} \qquad \text{such that} \ \begin{cases} D - B = 0 \\ D - C = 0 \end{cases}$$

Augmented Lagrangian formula:

$$L_\delta = f(B) + g(C) + h(D) + \langle Z_1, D - B\rangle + \langle Z_2, D - C\rangle +$$

$$\frac{\delta_1}{2}\|D - B\|_F^2 + \frac{\delta_2}{2}\|D - C\|_F^2.$$

Solution to $B^{[k+1]}$

Problem:

$$B^{[k+1]} := \underset{B}{\operatorname{argmin}} \left\{ 2f(B) \; + \; \delta_1^{[k]} \left\| D^{[k]} + \frac{Z_1^{[k]}}{\delta_1^{[k]}} - B \right\|_F^2 \right\}$$

Solution:

$$B^{[k+1]} = \frac{Y + \delta_1^k D^{[k]} + Z_1^k}{1 + \delta_1^k}$$

Solution to $C^{[k+1]}$

Problem:

$$C^{[k+1]} := \underset{C}{\operatorname{argmin}} \left\{ 2g(C) \ + \ \delta_2^{[k]} \left\| D^{[k]} + \frac{Z_2^{[k]}}{\delta_2^{[k]}} - C \right\|_F^2 \right\}$$

Solution:

$$\begin{cases} S^* := \operatorname{diag}\left(\left[\, (s_1^{[k]} - \frac{\lambda_N}{\delta_2^{[k]}})_+ \ , \ldots , \ (s_p^{[k]} - \frac{\lambda_N}{\delta_2^{[k]}})_+ \, \right]^T \right) \\ C^{[k+1]} = U^{[k]} S^* {V^{[k]}}^T \end{cases}$$

Solution to $D^{[k+1]}$

Problem:

$$D^{[k+1]} := \underset{D}{\operatorname{argmin}} \left\{ 2h(D) \ + \ \delta_1^{[k]} \left\| D - K \right\|_F^2 \ + \ \delta_2^{[k]} \left\| D - L \right\|_F^2 \right\}$$

Solution:

$$D_{ij}^{[k+1]} = \operatorname{sgn}\left(Q_{ij}^{[k+1]}\right) \cdot \left(\left| Q_{ij}^{[k+1]} \right| - \frac{\lambda_L W_{ij}}{\Delta^{[k]}} \right)_+ , \quad \text{for} \quad i, j \in \{1, \ldots, p\}$$