

Graphic Lasso: Data Analysis

Jixin Hu

February 14, 2021

1 Why the clustering results are not good?

1.1 More iterations

Last time, I consider the cut correlation matrices. See typical correlation pattern for three tissues in Figure 1, and corresponding generalized precision matrices in Figure 2.

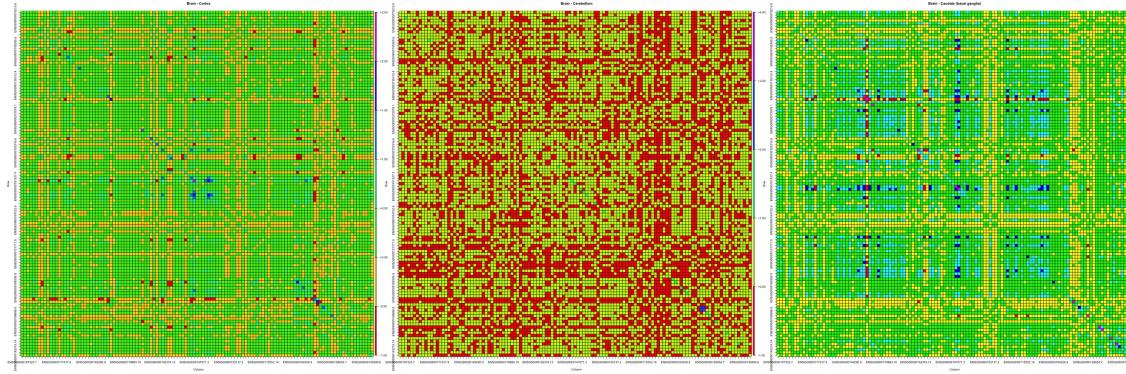


Figure 1: Sub-matrices for three brain tissues groups.

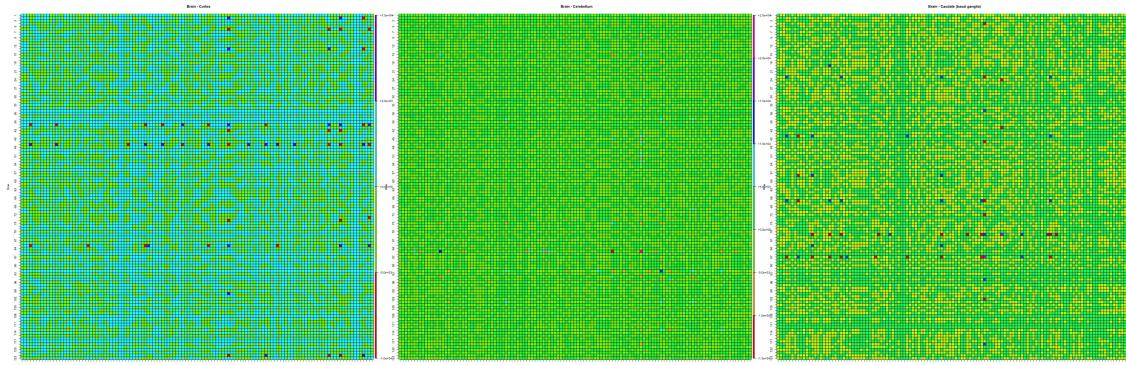


Figure 2: Precision matrices for three brain tissues groups.

The membership clustering results becomes better after more iterations. Below is the membership result after 250 iterations. See Table 1.

The algorithm is able to distinguish the cerebellum and other tissues. It seems like the previous bad results is caused by the sub-optimal solution. However, in A_2 , notice that Hippocampus is also

$A_0 : u_{ki} = 0$	Brain - Cortex Brain - Putamen (basal ganglia) Brain - Amygdala Brain - Hypothalamus Brain - Substantia nigra Brain - Anterior cingulate cortex (BA24) Brain - Frontal Cortex (BA9) Brain - Caudate (basal ganglia)
$A_1 : u_{k1} \neq 0$	Brain - Spinal cord (cervical c-1)
$A_2 : u_{k2} = 0$	Brain - Cerebellum Brain - Hippocampus Brain - Cerebellar Hemisphere
$A_3 : u_{k3} = 0$	Brain - Nucleus accumbens (basal ganglia)

Table 1: Membership result for 13 brain tissues after 250 iterations.

grouped with cerebellum. See Figure 3 for the correlation matrices. By the figure, Hippocampus should be grouped with cortex or basal ganglia.

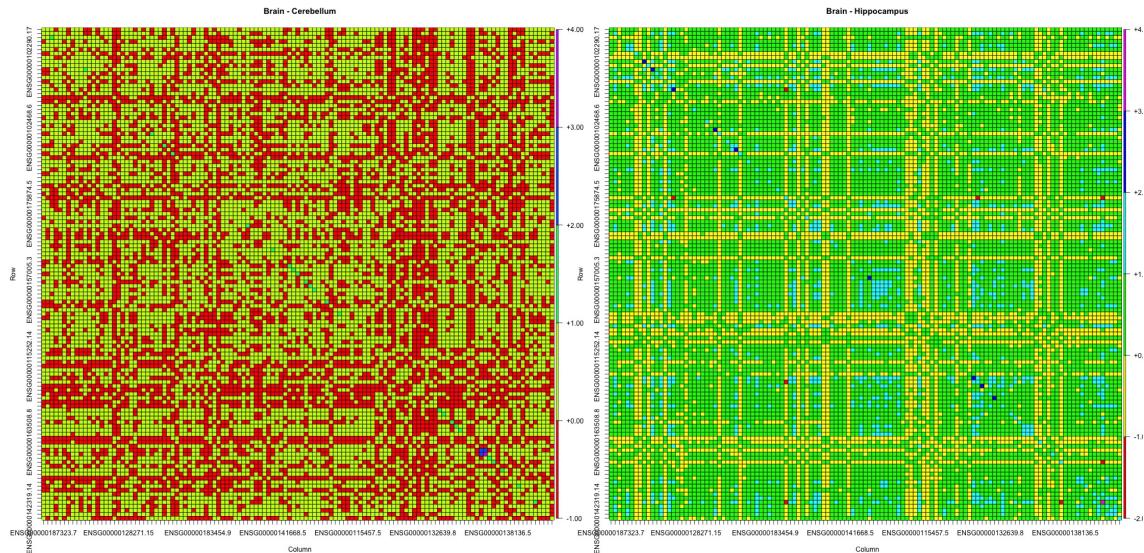


Figure 3: Correlation matrices for Cerebellum and Hippocampus.

Therefore, it needs to double check whether the optimality is the main reason for the bad clustering.

1.2 Optimality check

I select three tissues from Cortex-Basal ganglia group: Cortex, Putamen (basal ganglia), and Anterior cingulate cortex (BA24), and two tissues from cerebellum: Cerebellum, and Cerebellar Hemisphere.

I set $r = 1, \rho = 1000$. After 237 iterations, we obtain the membership which implies these five tissues share the same correlation structure with different magnitudes. The objective value is -70831 .

I use two ways to check the optimality.

1. Apply the Matlab algorithm to the cortex group and cerebellum group separately with $r = 1, \rho = 1000$. The convergent objective value of cortex group is -72801 ; the convergent objective value of cerebellum group is -32660 . Since the objective function is additive to

each tissue, the objective function for these 5 tissues is $-105461 < -70831$, which implies the algorithm does not achieve the optimal under the joint estimation.

2. Apply the single-layer Glasso to each tissue with $\rho = 0.005$. The sum of the negative log-likelihood is -69738 , which is at the same level of the joint estimation.

Based on the joint and separate performance with Matlab algorithm, the joint estimation seems give a sub-optimal result. The possible reason for single-layer Glasso performs worse than joint estimation may be that the Matlab algorithm does not apply sparse penalty to Θ_0 while Glasso can not handle the dense case very well.