Lab 6 - Advanced Machine Learning

Probabilistic graph model

Pierre HOUDOUIN

CentraleSupelec

03 Novembre 2023

General information

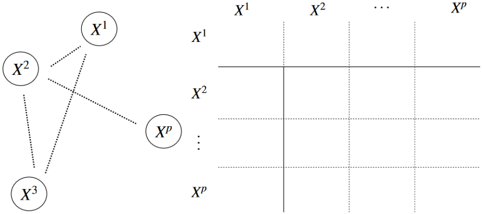
- **Assignment** : Alone or in pairs, you will code the algorithms you learnt in Ścikit-learn formalism'
- **Due** : Each assignment has to be sent at most 7 days after the lab session at pierre.houdouin@centralesupelec.fr
- **Grading** : There are 5 lab sessions, each lab session is worth 4 points, the average will count for half of your final grade
- **Questions** : If you have questions, comments or feedbacks about the lab session, feel free to contact me by email

Lesson recap

Graphical model

Suppose we have a random vector $X = (X_1, \ldots, X_p)$:

$$\begin{aligned} & \textit{no edge between node } i \textit{ and } j \\ \iff & X_i \textit{ and } X_j \textit{ are conditionally independant} \\ \iff & X_i \textit{ and } X_j \textit{ have } 0 \textit{ inverse correlation (gaussian case)} \end{aligned}$$



$\longrightarrow$ *need to estimate* $\Lambda = \Sigma^{-1}$

Estimation of the precision matrix

Classic

Use the Maximum Likelihood :

$$\min_{\Lambda > 0} -\log\det\Lambda + Tr(S\Lambda) \quad \underset{\text{closed-form}}{\Rightarrow} \quad \hat{\Lambda} = S^{-1} = \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)\left(x_i - \hat{\mu}\right)^T\right)^{-1} \quad \text{where } \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

inverse empirical covariance

empirical mean

not great when $n << p$: reduce #parameters via a sparsity constraint

GLASSO

Use the Sparse Maximum Likelihood :

$$\min_{\Lambda > 0} -\log\det\Lambda + Tr(S\Lambda) + \lambda\|\Lambda\|_1, \quad \lambda > 0 \quad \underset{\text{ADMM algo}}{\Rightarrow}$$

reformulate

$$\min_{\substack{\Lambda \\ Y = \Lambda}} -\log\det\Lambda + Tr(S\Lambda) + 1_{\Lambda > 0} + \lambda\|Y\|_1$$

$$T_k \leftarrow Y_{k-1} - Z_{k-1} - \frac{}{\nu}S, \quad \nu > 0$$

intermediate step for computations

weight of the penalization $Y = \Lambda$

$$\Lambda_k \leftarrow \frac{1}{2}\operatorname{diag}\left(\left\{\lambda_i(T_k) + \sqrt{\lambda_i(T_k)^2 + \frac{4}{\nu}}\right\}\right)$$

primal variable

$$Y_k \leftarrow \mathrm{ST}_{(\lambda/\nu)}\left(\Lambda_k + Z_{k-1}\right)$$

relaxed primal variable

$$Z_k \leftarrow Z_{k-1} + \nu\left(\Lambda_k - Y_k\right)$$

dual variable

Estimation of the precision matrix

Nodewise regression

Learn the support of the precision matrix via sparse linear regressions :

predict each variable X^j using the others $\{X^k, k \neq j\}$ with LASSO $\rightarrow \beta^j \in \mathbb{R}^{p-1}$ weights of the regression…

then deduce the support : $\Lambda_{jk} \neq 0 \iff \hat{\beta}^j_k \quad AND \quad \hat{\beta}^k_j$ are non-zero …used as edge-activators

(also try OR)

Assignment plan

- **Part 1** : Implement your own Graphical LASSO
- **Part 2** : Implement your own Node-wise Regression algorithm
- **Part 3** : Application to simulated data