# Assignment: Graphical LASSO

**Instructions**

- This assignment is **due on April 27**, 2025, at 23:59. It counts for 20% of the final grade.
- You can work **individually or in groups** of up to 3 people.
- Submit your **individual report** as a single PDF file on DTU Learn. The report should be **at most 6 pages**. Make sure to include the names and student numbers of all group members, if any, in the PDF.
- **Attach your code** as a zip file or a self-contained Jupyter notebook. Groups may write their code collaboratively. The use of AI is allowed but must disclosed and properly referenced. The code should be written in Python, Julia, or MATLAB.
- DTU's code of honour applies to this assignment.
- Questions about the assignment should be posted on the discussion forum on DTU Learn or asked during the exercise sessions.

## Problem formulation

Consider the "graphical LASSO" problem, which seeks a minimizer of a function $f : \mathbb{S}^n \to \bar{\mathbb{R}}$ of the form

$$f(X) = \operatorname{tr}(SX) - \ln\det X + \gamma \sum_{i \neq j} |X_{ij}|$$

with $\operatorname{dom} f = \mathbb{S}^n_{++}$ and where $S \in \mathbb{S}^n_{++}$ is a given covariance matrix, $\gamma \geq 0$ is a parameter, and $X \in \mathbb{S}^n$ is the variable. The sum is over all $i, j \in \mathbb{N}_n$ such that $i \neq j$, and $X$ represents the precision matrix of a Gaussian graphical model. LASSO is short for "least absolute shrinkage and selection operator" and is a popular approach to estimating sparse models in statistics and machine learning.

(1) Show that the function $f$ is convex, and derive the necessary and sufficient conditions for optimality.

*Hint:* To show that $-\ln\det(X)$ is convex on $\mathbb{S}^n_{++}$, it suffices to show that $-\ln\det(X + tV)$ is a convex function of $t$ for all $X \in \mathbb{S}^n_{++}$ and $V \in \mathbb{S}^n$ (*cf.* Exercise 4.5(b) in the textbook).

(2) For which values of $\gamma$ is $X^\star = \operatorname{diag}(S_{11}, \dots, S_{nn})^{-1}$ optimal?

(3) Show that the function $f$ is *coersive*, *i.e.*, show that if $\|X\| \to +\infty$ or $X$ approaches the boundary of $\operatorname{dom} f$ (which consists of singular positive semidefinite matrices), then $f(X) \to +\infty$. (This implies that the sublevel sets of $f$ are compact.)

(4) Consider the equivalent formulation of the graphical LASSO problem

$$\begin{array}{ll} \text{minimize} & g(X) + h(Y) \\ \text{subject to} & X = Y \end{array}$$

where $g : \mathbb{S}^n \to \bar{\mathbb{R}}$ is defined as $g(X) = \operatorname{tr}(SX) - \ln\det(X)$ with $\operatorname{dom} g = \mathbb{S}^n_{++}$ and $h : \mathbb{S}^n \to \mathbb{R}$ is defined as $h(Y) = \gamma \sum_{i \neq j} |Y_{ij}|$.

Show that the dual of this reformulation can be formulated as

$$\begin{array}{ll} \text{maximize} & \ln\det(S + U) + n \\ \text{subject to} & |U_{ij}| \leq \gamma, \quad i \neq j \\ & U_{ii} = 0, \quad i \in \mathbb{N}_n \end{array}$$

with variable $U \in \mathbb{S}^n$, and show that $X^\star = (S + U^\star)^{-1}$ is optimal if $U^\star$ is a dual optimal solution.

(5) Implement a function that takes $S \in \mathbb{S}^n_{++}$ and $\gamma > 0$ and solves the graphical LASSO problem using CVXPY (Python), Convex.jl (Julia), or CVX (MATLAB).

(6) Let $S$ be given by

$$S = \begin{bmatrix} 1.00 & -0.50 & 0.10 & -0.90 \\ -0.50 & 1.25 & -0.05 & 1.05 \\ 0.10 & -0.05 & 0.26 & -0.09 \\ -0.90 & 1.05 & -0.09 & 5.17 \end{bmatrix}$$

and let $X^\star(\gamma)$ be the solution of the graphical LASSO problem as a function of $\gamma$. The points

$$\begin{bmatrix} \operatorname{tr}(SX^\star(\gamma) - \ln\det(X^\star(\gamma)) \\ \sum_{i \neq j} |X^\star_{ij}(\gamma)| \end{bmatrix}$$

define a *trade-off curve*, which is parameterized by $\gamma$. Compute points on this curve for different values of $\gamma$ ranging from $10^{-2}$ to $10^{0}$ using your code from the previous question, and plot these points in a 2D plot.

*Hint:* Use the function `logspace` (NumPy/MATLAB) or `exp10` (Julia) to generate, say, 10 values of $\gamma$ that are evenly spaced on a logarithmic scale.

(7) Implement a proximal gradient method for solving the graphical LASSO problem with

$$f(X) = g(X) + h(X)$$

where $g(X) = \mathrm{tr}(SX) - \ln\det(X)$ and $h(X) = \gamma \sum_{i \neq j} |X_{ij}|$.

- Use a backtracking line search to find a step size that satisfies

$$X_{k+1} \succ 0, \quad g(X_{k+1}) \leq g(X_k) + \langle \nabla g(X_k), X_{k+1} - X_k \rangle + \frac{1}{2t_k} \|X_{k+1} - X_k\|_{\mathrm{F}}^2,$$

  where

$$X_{k+1} = \mathrm{prox}_{t_k h}(X_k - t_k \nabla g(X_k)).$$

- Use a stopping criterion of the form $\Delta \leq \epsilon$ where $\epsilon > 0$ is a small positive number (say, $\epsilon = 10^{-2}$) and where

$$\Delta = g(X) + h(X) - \ln\det(S + U) - n$$

  is the duality gap at a primal-dual pair $(X, U)$. Given the primal variable $X \in \mathbb{S}^n_{++}$, we can choose $U \in \mathbb{S}^n$ as

$$U = \mathrm{prox}_{h^*}(X^{-1} - S),$$

  which is the matrix with elements

$$U_{ij} = \begin{cases} \max(-\gamma, \min(\gamma, [X^{-1} - S]_{ij})), & i \neq j, \\ 0, & i = j. \end{cases}$$

  This is motivated by the optimality condition $U = X^{-1} - S$ and the constraints in the dual problem. Note that $S + U$ is not necessarily positive definite with this choice, in which case the duality gap at $(X, U)$ is not finite.

(8) Download the data file `sp500.txt`, which contains a covariance matrix computed from data from the S&P 500 index. Use this covariance matrix as the matrix $S$ and solve the graphical LASSO problem for different values of $\gamma$ using your implementation of the proximal gradient method. Use 10 values of $\gamma$ that are evenly spaced on a logarithmic scale and ranging from $10^{-2}$ to $10^{-1}$.

- Plot a trade-off curve for this problem as in question (5).
- Plot the number of off-diagonal nonzeros in $X^\star(\gamma)$ as a function of $\gamma$.

*Hint:* You may use the solution to the graphical LASSO problem with a given $\gamma$ as an initial guess when solving with a different (but nearby) value of $\gamma$.