General remarks:

- You are allowed to use two A4 pages (either one A4 sheet front and back or two A4 sheets front only) of machine- or handwritten notes.
- Pocket calculators are not allowed.
- Switch off your mobile phone and any other electronic device with communication capabilities. These are not allowed to be on the table.
- Questions/Issues/Toilet: Please raise your hand and we will come to you.
- There is exactly one correct answer to each question, namely either "True" or "False".
- A **correct** answer (which might be "True" or "False") will give **1 point**, an **incorrect answer 0 point** and **no answer 0 points**.
  Hence: you should always make at least a guess even when being highly unsure about which answer would be true.
- Each answer has to be clearly marked with a black or a dark blue pen (no pencil). Two or more marks in a question or unclear marks give zero points for the question.
- We only grade the "True"/"False" marks on the **answer sheet**. Texts, numbers, and other solutions will be ignored.

Good Luck!

Lasso for linear models.
Consider a high-dimensional linear model with a continuous response $Y \in \mathbb{R}$ and p real-valued fixed (deterministic) covariates $X^{(1)}, \ldots, X^{(p)} \in \mathbb{R}$ that are connected through the model

$$Y_i = \sum_{j=1}^{p} \beta_j^0 X_i^{(j)} + \varepsilon_i \quad (i = 1, \ldots, n)$$

with $p \gg n$ and where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. with $\mathbb{E}[\varepsilon_1] = 0$. In matrix notation, we can equivalently write

$$Y = X\beta^0 + \varepsilon.$$

The $n \times p$ design matrix X is fixed (deterministic).

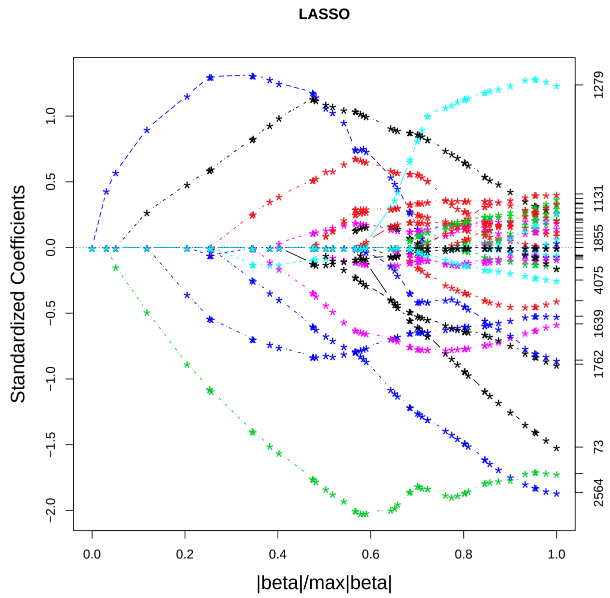
1. In this setting (as described in the introductory prompt of this section), assess the following statements:

   a) An ordinary least squares (OLS) estimator exists but it is not unique and can be expected to overfit the data.

   b) Assume that $\text{rank}(X) = n$ (which is a reasonable assumption if $p \gg n$). Then, an ordinary least squares (OLS) estimator exists and has residual sum of squares equal to zero.

2. For the following questions in this problem, we drop the assumption that $p \gg n$ necessarily holds.

   Consider the following Lasso regularization path (for a certain dataset).



   Assess the following statements:

   a) When decreasing $\lambda \searrow 0$ one obtains decreasing values on the x-axis.
   b) Lasso regularization paths are always monotone, that is, if the coefficient estimate is non-zero for some value x_0 on the x-axis it remains non-zero for all values $x > x_0$.
   c) Lasso regularization paths never cross the x-axis, that is, they cannot first be positive and then negative or the other way around.

3. Assess the following statement: The Ridge estimator for a fixed regularization parameter can be computed with an explicit algebraic formula (without the need to use a convex optimization algorithm).

4. Assume that $\lambda = C\sqrt{\log(p)/n}$ with $0 < C < \infty$ a sufficiently large constant. Furthermore, we assume $s_0 = ||\beta^0||_1 = o(\sqrt{n/\log(p)})$ and assume Gaussian errors $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T \sim \mathcal{N}(0, \sigma^2 I_{n\times n})$ for a strictly positive real number σ^2 and where the symbol $I_{n\times n}$ denotes the $n \times n$ identity matrix). Consider the Lasso estimator $\hat{\beta}(\lambda)$.

   Assess the following statements:

   a) For any design X where $\frac{1}{n}\sum_{i=1}^{n}(X_i^{(j)})^2 = 1$ holds for all $j = 1, \dots, p$, the Lasso is consistent for prediction with respect to the ℓ_2-norm, meaning that

   $$\|X(\hat{\beta}(\lambda) - \beta^0)\|_2^2/n \to 0 \ (p \geq n \to \infty) \text{ in probability.}$$

   b) For any design X where $\frac{1}{n}\sum_{i=1}^{n}(X_i^{(j)})^2 = 1$ holds for all $j = 1, \dots, p$, the Lasso is consistent for parameter estimation with respect to the ℓ_1-norm, meaning that

   $$\|\hat{\beta}(\lambda) - \beta^0\|_1 \to 0 \ (p \geq n \to \infty) \text{ in probability.}$$

Consider a high-dimensional linear model with a continuous response $Y \in \mathbb{R}$ and p real-valued fixed (deterministic) covariates $X^{(1)}, \ldots, X^{(p)} \in \mathbb{R}$ that are connected through the model

$$Y_i = \sum_{j=1}^{p} \beta_j^0 X_i^{(j)} + \varepsilon_i \quad (i = 1, \ldots, n)$$

with $p \gg n$ and where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. with $\varepsilon_1 \sim \mathcal{N}(0, \sigma^2)$. In matrix notation, we can equivalently write

$$Y = X\beta^0 + \varepsilon.$$

The $n \times p$ design matrix X is fixed (deterministic).

5. Assess the following statement: The de-biased Lasso estimate $\hat{b}$ is a sparse estimator in the sense that there is a range of values for the regularization parameter(s) such that the corresponding de-biased Lasso estimate satisfies: there exists $j \in \{1, \ldots, p\}$ with $\hat{b}_j = 0$.

6. Consider a Gaussian graphical model where $(X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)}, X^{(5)}) \sim \mathcal{N}_5(0, \Sigma)$ with

$$\Sigma = \begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix} \quad \text{and} \quad \Sigma^{-1} = \begin{pmatrix} 1.2 & -0.6 & 0.0 & 0.8 & 0.4 \\ -0.6 & 0.8 & 0.0 & -0.4 & -0.2 \\ 0.0 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.8 & -0.4 & 0.0 & 1.2 & 0.6 \\ 0.4 & -0.2 & 0.0 & 0.6 & 0.8 \end{pmatrix}.$$

Assess the following statement: There is an undirected edge between the nodes X_2 and X_4.

7. Assess the following statement: The Graphical Lasso (GLasso) can be expressed as regularized maximum likelihood estimation with regularization penalty being equal to $\lambda \sum_{j=1}^{p} \sum_{k=1}^{p} |\Sigma_{j,k}|$ where Σ is the unknown covariance matrix.

Solutions:

1a) TRUE

1b) TRUE

2a) FALSE

2b) FALSE

2c) FALSE

3 TRUE

4a) TRUE

4b) FALSE

5 FALSE

6 TRUE

7 FALSE