Notes. **NO** late submission will be accepted. Computer generated output without detailed explanations and remarks will not receive any credit. You may type out your answers, but make sure to use different fonts to distinguish your own words from computer output. Only hard copies are accepted. For the simulation and data analysis problems, keep the code you develop as you may be asked to present your work in class.

**1.** Consider the statement regrading (3.7) on P.81 of CO.

(a) Formulate the corresponding version precisely for concave functions. In particular, you should clearly specify the domains of all involved functions.

(b) Use Part (a) to show that the dual function is always concave.

**2.** Show whether strong duality holds for ridge regression and Lasso when $t = 0$.

**3.** Show that the set $\mathcal{A}$ defined in (5.37) on P.232 of CO is convex, if the underlying optimization problem is convex.

**4.** Exercise 5.1 (a), (b), (c) of CO.

**5.** Exercise 5.14 of CO.

**6.** Exercise 5.27 of CO, assuming that the primal problem is feasible, i.e. $Gx = h$ has solutions.

**7.** Assume $\boldsymbol{X} \in \mathbb{R}^{N\times p}$ is full rank, i.e. $\operatorname{rank}(\boldsymbol{X}) = p$, and assume $\boldsymbol{y} \neq \mathbf{0}$. Consider the two formulations of Lasso. The first one is

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 \\ \text{subject to} & \|\boldsymbol{\beta}\|_1 \leq t. \end{array} \tag{P}$$

Denote by $\hat{\boldsymbol{\beta}}(t)$ the optimal solution of (P). The other formulation is

$$\min_{\boldsymbol{\beta}} \left\{ \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + \lambda\|\boldsymbol{\beta}\|_1 \right\}. \tag{D}$$

Denote by $\hat{\boldsymbol{\beta}}_\lambda$ the optimal solution of (D). Let $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ be the LSE. Prove that for each $0 < t < \|\hat{\boldsymbol{\beta}}\|_1$, there exists a $\lambda > 0$, such that $\hat{\boldsymbol{\beta}}_\lambda = \hat{\boldsymbol{\beta}}(t)$.