Convex Optimization: Homework 3

Alexandre ver Hulst

December 2nd, 2024

1 Question 1

Answer: We have

$$\begin{aligned} \text{LASSO} &\iff \min_{\mathbf{w}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{Xw}-\mathbf{y}\|_2^2 + \lambda\|\mathbf{w}\|_1 \\ &\iff \min_{\mathbf{w}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{Xw}-\mathbf{y}\|_2^2 + \lambda\|\mathbf{w}\|_1 \quad \text{s.t.} \quad \mathbf{z} = \mathbf{Xw}-\mathbf{y} \end{aligned}$$

We now consider the Lagrangian to compute the dual problem. We thus have

$$L(\mathbf{w},\mathbf{z},\nu) = \frac{1}{2}\|\mathbf{z}\|_2 + \lambda\|\mathbf{w}\|_1 + \nu^T(\mathbf{z}-\mathbf{Xw}-\mathbf{y})$$

We want

$$\inf_{\mathbf{w},\mathbf{z}\in\mathbb{R}^n} L(\mathbf{w},\mathbf{z},\nu) = \inf_{\mathbf{z}\in\mathbb{R}^n} [\frac{1}{2}\|\mathbf{z}\|_2^2 + \nu^{\mathbf{T}}\mathbf{z}] + \inf_{\mathbf{w}\in\mathbb{R}^n} [\lambda\|\mathbf{w}\|_1 - \nu^{\mathbf{T}}\mathbf{Xw}] + \nu^{\mathbf{T}}\mathbf{y}$$

We consider $g(\mathbf{z}) = \frac{1}{2}\|\mathbf{z}\|_2^2 + \nu^{\mathbf{T}}\mathbf{y}$ We have $\nabla g(\mathbf{z}) = \mathbf{z} + \nu = 0 \iff \nu = -\mathbf{z}$. In this case, as g is convex (as sum of the norm, which is convex, and a linear function which is evidently convex) and differentiable (as sum on two such functions), g is minimized at the critical point $\mathbf{z} = -\nu$, and we have $\inf_{\mathbf{z}} g(\mathbf{z}) = -\frac{1}{2}\|\nu\|_2^2$. On the other hand, we consider $f(\mathbf{w}) = \|\mathbf{w}\|_1$. We have the conjugate $f^*(y) = \sup_{\mathbf{w}\in\mathbb{R}^n} \mathbf{y}^T\mathbf{w} - \|\mathbf{w}\|_1$. As this conjugate function has already been calculated in Homework 2, we will directly use the result without redoing it. We thus have $\lambda f^*(\frac{1}{\lambda}\mathbf{X}^T\nu) = \begin{cases} 0 & \text{if } \|\mathbf{X}^T\nu\|_\infty \leq \lambda \\ -\infty & \text{otherwise} \end{cases}$ We thus have

$$\begin{aligned} \text{LASSO} &\iff \sup_{\nu\in\mathbb{R}^n} -\frac{1}{2}\|\nu\|_2^2 + \nu^T\mathbf{y} \quad \text{s.t.} \quad \|\mathbf{X}^T\nu\|_\infty \leq \lambda \\ &\iff \sup_{\nu\in\mathbb{R}^n} \nu^T(-\frac{1}{2}\mathbf{I}_n)\nu + \mathbf{y}^T\nu \quad \text{s.t.} \quad \begin{pmatrix} X^T \\ -X^T \end{pmatrix}\nu \leq \lambda\mathbf{1}_{2d} \\ &\iff \inf_{\nu\in\mathbb{R}^n} \nu^T(\frac{1}{2}\mathbf{I}_n)\nu - \mathbf{y}^T\nu \quad \text{s.t.} \quad \begin{pmatrix} X^T \\ -X^T \end{pmatrix}\nu \leq \lambda\mathbf{1}_{2d} \\ &\iff \inf_{\mathbf{v}\in\mathbb{R}^n} \mathbf{v}^T\mathbf{Q}\nu + \mathbf{p}^T\mathbf{v} \quad \text{s.t.} \quad \mathbf{Av} \leq \mathbf{b} \text{ where} \begin{cases} \mathbf{Q} = \frac{1}{2}\mathbf{I}_n \\ \mathbf{p} = -\mathbf{y} \\ \mathbf{A} = \begin{pmatrix} X^T \\ -X^T \end{pmatrix} \\ \mathbf{b} = \lambda\mathbf{1}_{2d} \end{cases} \end{aligned}$$

2 Question 2

Answer: See Notebook

3 Question 3

Answer: See Notebook