Optimization Theory and Algorithm II	**September 15, 2022**

Homework 1: Review

Lecturer:Xiangyu Chang *Scribe: Xiangyu Chang*

Edited by: Xiangyu Chang

HW 1 *In the Portfolio Management example, we formulate the oprimization problem as:*

$$\max_{\mathbf{x}} \ \mathbb{E}(R) - \lambda Var(R), \tag{1}$$
$$s.t. \ x_i \geq 0, i = 1 \ldots, n, \tag{2}$$
$$\sum_{i=1}^{n} x_i = 1, \tag{3}$$

where $\mathbb{E}(R)$ is the expectation of R, $Var(R)$ is the variance of R and $\lambda > 0$ is called risk aversion parameter for balancing the investment risk and expected return. We further suppose that $\mathbf{r} = (r_1, \ldots, r_n)^\top$ is a random vector with expectation of $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$, and covariance of Σ. Then prove that the optimization problem (1) is equivalent to

$$\max_{\mathbf{x}} \ \boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \Sigma \mathbf{x}, \tag{4}$$
$$s.t. \ x_i \geq 0, i = 1 \ldots, n, \tag{5}$$
$$\sum_{i=1}^{n} x_i = 1. \tag{6}$$

HW 2 *Poisson Regression: Suppose now that the b_i are counts and the $\mathbf{a}_i$ are some predictors in $\mathbb{R}^p$. Let λ_i be the expected number of events in sometime period, i.e.*

$$P(b_i|\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{b_i}}{b_i!}.$$

Based on the example of Generalized Linear Model, please derive the following optimization form to present the simple Poisson regression.

$$\min_{\mathbf{x}} \ \sum_{i=1}^{m} \exp(\langle \mathbf{a}_i, \mathbf{x} \rangle) - \langle \mathbf{b}, A\mathbf{x} \rangle.$$

HW 3 *Let us consider the LASSO problem:*

$$\min_{\mathbf{x}} \{ \frac{1}{2} \|A\mathbf{x} - \mathbf{b}\|^2 + \lambda \|\mathbf{x}\|_1 \}. \tag{7}$$

- *If $\lambda = 0$, provide the gradient descent algorithm (step size is related to β-smooth parameter.*
- *If $\lambda = 0$, provide the Newton-Raphson algorithm. Justify one-step Newton-Raphson algorithm can find the optimal solution.*
- *If $\lambda > 0$, provide the sub-gradient descent algorithm.*
- *If $\lambda > 0$, provide the proximal gradient descent algorithm.*

Note: show the detailed derivations.

HW 4 *Consider the following convex optimization problem:*

$$\min_{\mathbf{x}} f(\mathbf{x}),$$
$$s.t.\ \mathbf{x} \preceq 0.$$

Then, according to general optimality condition prove that the optimal point $\mathbf{x}^*$ *satisfies*

$$\nabla f(\mathbf{x}^*) \preceq 0 \text{ and } x_i^*(\nabla f(\mathbf{x}^*))_i = 0, i = 1, \ldots, n.$$

HW 5 *First, recall the support vector machine in the lecture notes. Then consider the following relaxed linear SVM model*

$$\min_{\mathbf{w},\mathbf{b},\xi} \frac{\|\mathbf{w}\|^2}{2} + C\sum_{i=1}^{n} \xi_i$$
$$s.t.\ y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + \mathbf{b}) \geq 1 - \xi_i,$$
$$\xi_i \geq 0, i = 1, \ldots, n,$$

where C *is a constant.*

Please show its Lagrange dual problem.

References