ELE 535 - Machine Learning and Pattern Recognition

Fall 2018

HOMEWORK 4

Q1 Determine general sufficient conditions (if any exist) under which the indicated function f is convex.
(a) $f\colon \mathbb{R} \to \mathbb{R}$ with $f(x) = |x|$.
(b) $f\colon (0, \infty) \to \mathbb{R}$ with $f(x) = x \ln(x)$.
(c) $f\colon \mathbb{R}^n \to \mathbb{R}$ with $f(x) = (x^T Q x)^3$. Here $Q \in \mathbb{R}^{n\times n}$ is symmetric PSD.
(d) $f\colon \mathbb{R}^n \to \mathbb{R}$ with $f(x) = 1 + e^{(\sum_{i=1}^n |x(i)|)^3}$.
(e) For $x \in \mathcal{C} = \{x \in \mathbb{R}^n \colon x(i) > 0, i = 1, \dots, n\}$ let $\ln(x) = [\ln(x(i))] \in \mathbb{R}^n$ and define $f(x) = x^T \ln(x)$.

Q2 You want to learn an unknown function $f\colon [0,1] \to \mathbb{R}$ using a set of noisy measurements (x_j, y_j), with $y_j = f(x_j) + \epsilon_j$, $j = 1, \dots, m$. Your plan is to approximate $f(\cdot)$ by a Fourier series on $[0,1]$ with $q \in \mathbb{N}$ terms:

$$f_q(x) = \frac{a_0}{2} + \sum_{k=1}^{q} a_k \cos(2\pi k x) + b_k \sin(2\pi k x).$$

To control the smoothness of $f_q(\cdot)$, you also decide to penalize the size of the coefficients a_k, b_k more heavily as k increases.

(a) Formulate the above problem as a regularized regression problem.

(b) For $q = 2$, display the regression matrix, the label vector y, and the regularization term.

(c) Comment briefly on how to select q.

Q3 Let $D \in \mathbb{R}^{n\times n}$ be diagonal with nonnegative diagonal entries and consider the problem:

$$\min_{x\in\mathbb{R}^n} \quad \|x - y\|_2^2 + \lambda \|Dx\|_2^2.$$

This problem seeks to best approximate $y \in \mathbb{R}^n$ with a nonuniform penalty for a large entries in x.
(a) Solve this problem using the solution of ridge regression.
(b) Show that the objective function is separable into a sum of decoupled terms. Show that this decomposes the problem into n independent scalar problems.
(c) Find the solution of each scalar problem.
(d) By putting these scalar solutions together, find and interpret the solution to the original problem.

Q4 Let $X \in \mathbb{R}^{n\times m}$ and $y \in \mathbb{R}^m$ be given, and $\lambda > 0$. Consider the problem

$$w^\star = \arg\min_{w\in\mathbb{R}^n} \|y - X^T w\|_2^2 + \lambda \|w\|_2^2. \tag{1}$$

From the notes we that there exists a unique solution $w^\star$ and that $w^\star \in \mathcal{R}(X)$. Using (1) and these two results, show that $w^\star = X(X^T X + \lambda I_m)^{-1} y$.

Q5 One form of regularized least squares can be posed as:

$$\min_{w\in\mathbb{R}^n} \quad \|Fw - y\|_2^2 + \lambda \|Gw - g\|_2^2, \tag{2}$$

where $F \in \mathbb{R}^{m\times n}$, $y \in \mathbb{R}^m$, $G \in \mathbb{R}^{k\times n}$, $g \in \mathbb{R}^k$ and $\lambda > 0$.
(a) Show that a sufficient condition for (2) to have a unique solution is that $\operatorname{rank}(G) = n$.
(b) Show that a necessary and sufficient condition is that $\mathcal{N}(F) \cap \mathcal{N}(G) = \mathbf{0}$.