HW V

EE 588: Optimization for the information and data sciences

University of Southern California

Assigned on: October 24, 2018 Due date: beginning of class on November 7, 2018

---

1. *Estimating a vector with unknown measurement nonlinearity.* (A specific instance of exercise 7.9 in *Convex Optimization*.) We want to estimate a vector $x \in \mathbb{R}^n$, given some measurements
$$y_i = \phi(a_i^T x + v_i), \quad i = 1, \ldots, m.$$
Here $a_i \in \mathbb{R}^n$ are known, v_i are IID $\mathcal{N}(0, \sigma^2)$ random noises, and $\phi : \mathbb{R} \to \mathbb{R}$ is an unknown monotonic increasing function, known to satisfy
$$\alpha \le \phi'(u) \le \beta,$$
for all u. (Here α and β are known positive constants, with $\alpha < \beta$.) We want to find a maximum likelihood estimate of x and ϕ, given y_i. (We also know a_i, σ, α, and β.)

This sounds like an infinite-dimensional problem, since one of the parameters we are estimating is a function. In fact, we only need to know the m numbers $z_i = \phi^{-1}(y_i)$, $i = 1, \ldots, m$. So by estimating ϕ we really mean estimating the m numbers $z_1, \ldots, z_m$. (These numbers are not arbitrary; they must be consistent with the prior information $\alpha \le \phi'(u) \le \beta$ for all u.)

   (a) Explain how to find a maximum likelihood estimate of x and ϕ (i.e., $z_1, \ldots, z_m$) using convex optimization.

   (b) Carry out your method on the data given in `nonlin_meas_data.m`, which includes a matrix $A \in^{m \times n}$, with rows $a_1^T, \ldots, a_m^T$. Give $\hat{x}_{\text{ml}}$, the maximum likelihood estimate of x. Plot your estimated function $\hat{\phi}_{\text{ml}}$. (You can do this by plotting $(\hat{z}_{\text{ml}})_i$ versus y_i, with y_i on the vertical axis and $(\hat{z}_{\text{ml}})_i$ on the horizontal axis.)

   Hint. You can assume the measurements are numbered so that y_i are sorted in nondecreasing order, i.e., $y_1 \le y_2 \le \cdots \le y_m$. (The data given in the problem instance for part (b) is given in this order.)

2. *Maximum likelihood estimation of an increasing nonnegative signal.* We wish to estimate a scalar signal $x(t)$, for $t = 1, 2, \ldots, N$, which is known to be nonnegative and monotonically nondecreasing:
$$0 \le x(1) \le x(2) \le \cdots \le x(N).$$
This occurs in many practical problems. For example, $x(t)$ might be a measure of wear or deterioration, that can only get worse, or stay the same, as time t increases. We are also given that $x(t) = 0$ for $t \le 0$.

We are given a noise-corrupted moving average of x, given by

$$y(t) = \sum_{\tau=1}^{k} h(\tau)x(t-\tau) + v(t), \quad t = 2, \ldots, N+1,$$

where $v(t)$ are independent $\mathcal{N}(0,1)$ random variables.

(a) Show how to formulate the problem of finding the maximum likelihood estimate of x, given y, taking into account the prior assumption that x is nonnegative and monotonically nondecreasing, as a convex optimization problem. Be sure to indicate what the problem variables are, and what the problem data are.

(b) We now consider a specific instance of the problem, with problem data (i.e., N, k, h, and y) given in the file `ml_estim_incr_signal_data.m`. (This file contains the true signal `xtrue`, which of course you cannot use in creating your estimate.) Find the maximum likelihood estimate $\hat{x}_{\rm ml}$, and plot it, along with the true signal. Also find and plot the maximum likelihood estimate $\hat{x}_{\rm ml,free}$ *not taking into account the signal nonnegativity and monotonicity.*

Hint. The function `conv` (convolution) is overloaded to work with CVX.

3. *Three-way linear classification.* We are given data

$$x^{(1)}, \ldots, x^{(N)}, \quad y^{(1)}, \ldots, y^{(M)}, \quad z^{(1)}, \ldots, z^{(P)},$$

three nonempty sets of vectors in ^{n}. We wish to find three affine functions on $\mathbb{R}^n$,

$$f_i(z) = a_i^T z - b_i, \quad i = 1, 2, 3,$$

that satisfy the following properties:

$$\begin{array}{ll} f_1(x^{(j)}) > \max\{f_2(x^{(j)}), f_3(x^{(j)})\}, & j = 1, \ldots, N, \\ f_2(y^{(j)}) > \max\{f_1(y^{(j)}), f_3(y^{(j)})\}, & j = 1, \ldots, M, \\ f_3(z^{(j)}) > \max\{f_1(z^{(j)}), f_2(z^{(j)})\}, & j = 1, \ldots, P. \end{array}$$

In words: f_1 is the largest of the three functions on the x data points, f_2 is the largest of the three functions on the y data points, f_3 is the largest of the three functions on the z data points. We can give a simple geometric interpretation: The functions f_1, f_2, and f_3 partition $\mathbb{R}^n$ into three regions,

$$\begin{aligned} R_1 &= \{z \mid f_1(z) > \max\{f_2(z), f_3(z)\}\}, \\ R_2 &= \{z \mid f_2(z) > \max\{f_1(z), f_3(z)\}\}, \\ R_3 &= \{z \mid f_3(z) > \max\{f_1(z), f_2(z)\}\}, \end{aligned}$$

defined by where each function is the largest of the three. Our goal is to find functions with $x^{(j)} \in R_1$, $y^{(j)} \in R_2$, and $z^{(j)} \in R_3$.

Pose this as a convex optimization problem. You may not use strict inequalities in your formulation.

Solve the specific instance of the 3-way separation problem given in `sep3way_data.m`, with the columns of the matrices `X`, `Y` and `Z` giving the $x^{(j)},\ j = 1, \ldots, N$, $y^{(j)},\ j = 1, \ldots, M$ and $z^{(j)},\ j = 1, \ldots, P$. To save you the trouble of plotting data points and separation boundaries, we have included the plotting code in `sep3way_data.m`. (Note that `a1`, `a2`, `a3`, `b1` and `b2` contain arbitrary numbers; you should compute the correct values using CVX.)