# EE364a Homework 6 additional problems

1. *Minimax rational fit to the exponential.* (See exercise 6.9.) We consider the specific problem instance with data

$$t_i = -3 + 6(i-1)/(k-1), \quad y_i = e^{t_i}, \quad i = 1, \ldots, k,$$

where $k = 201$. (In other words, the data are obtained by uniformly sampling the exponential function over the interval $[-3, 3]$.) Find a function of the form

$$f(t) = \frac{a_0 + a_1 t + a_2 t^2}{1 + b_1 t + b_2 t^2}$$

that minimizes $\max_{i=1,\ldots,k} |f(t_i) - y_i|$. (We require that $1 + b_1 t_i + b_2 t_i^2 > 0$ for $i = 1, \ldots, k$.)

Find optimal values of $a_0, a_1, a_2, b_1, b_2$, and give the optimal objective value, computed to an accuracy of 0.001. Plot the data and the optimal rational function fit on the same plot. On a different plot, give the fitting error, *i.e.*, $f(t_i) - y_i$.

*Hint.* You can use `strcmp(cvx_status,'Solved')`, after `cvx_end`, to check if a feasibility problem is feasible.

2. *Maximum likelihood prediction of team ability.* A set of $n$ teams compete in a tournament. We model each team's ability by a number $a_j \in [0, 1]$, $j = 1, \ldots, n$. When teams $j$ and $k$ play each other, the probability that team $j$ wins is equal to $\textbf{prob}(a_j - a_k + v > 0)$, where $v \sim \mathcal{N}(0, \sigma^2)$.

You are given the outcome of $m$ past games. These are organized as

$$(j^{(i)}, k^{(i)}, y^{(i)}), \quad i = 1, \ldots, m,$$

meaning that game $i$ was played between teams $j^{(i)}$ and $k^{(i)}$; $y^{(i)} = 1$ means that team $j^{(i)}$ won, while $y^{(i)} = -1$ means that team $k^{(i)}$ won. (We assume there are no ties.)

(a) Formulate the problem of finding the maximum likelihood estimate of team abilities, $\hat{a} \in \mathbf{R}^n$, given the outcomes, as a convex optimization problem. You will find the *game incidence matrix* $A \in \mathbf{R}^{m \times n}$, defined as

$$A_{il} = \begin{cases} y^{(i)} & l = j^{(i)} \\ -y^{(i)} & l = k^{(i)} \\ 0 & \text{otherwise,} \end{cases}$$

useful.

The prior constraints $\hat{a}_i \in [0, 1]$ should be included in the problem formulation. Also, we note that if a constant is added to all team abilities, there is no change in the probabilities of game outcomes. This means that $\hat{a}$ is determined only up to a constant, like a potential. But this doesn't affect the ML estimation problem, or any subsequent predictions made using the estimated parameters.

(b) Find $\hat{a}$ for the team data given in `team_data.m`, in the matrix `train`. (This matrix gives the outcomes for a tournament in which each team plays each other team once.)

CVX does not support the concave function $\log \Phi$, where $\Phi$ is the cumulative distribution of a unit Gaussian, but we have provided a good enough approximation, `log_normcdf`, on the course web site. This function is overloaded to handle vector inputs (elementwise).

You can form $A$ using the commands

```
A = sparse(1:m,train(:,1),train(:,3),m,n) + ...
    sparse(1:m,train(:,2),-train(:,3),m,n);
```

(c) Use the maximum likelihood estimate $\hat{a}$ found in part (b) to predict the outcomes of next year's tournament games, given in the matrix `test`, using $\hat{y}^{(i)} = \mathbf{sign}(\hat{a}_{j^{(i)}} - \hat{a}_{k^{(i)}})$. Compare these predictions with the actual outcomes, given in the third column of `test`. Given the fraction of correctly predicted outcomes.

The games played in `train` and `test` are the same, so another, simpler method for predicting the outcomes in `test` it to just assume the team that won last year's match will also win this year's match. Give the percentage of correctly predicted outcomes using this simple method.

3. *Piecewise-linear fitting.* In many applications some function in the model is not given by a formula, but instead as tabulated data. The tabulated data could come from empirical measurements, historical data, numerically evaluating some complex expression or solving some problem, for a set of values of the argument. For use in a convex optimization model, we then have to fit these data with a convex function that is compatible with the solver or other system that we use. In this problem we explore a very simple problem of this general type.

Suppose we are given the data $(x_i, y_i)$, $i = 1, \ldots, m$, with $x_i,\ y_i \in \mathbf{R}$. We will assume that $x_i$ are sorted, *i.e.*, $x_1 < x_2 < \cdots < x_m$. Let $a_0 < a_1 < a_2 < \cdots < a_K$ be a set of fixed knot points, with $a_0 \leq x_1$ and $a_K \geq x_m$. Explain how to find the convex piecewise linear function $f$, defined over $[a_0, a_K]$, with knot points $a_i$, that minimizes the least-squares fitting criterion

$$\sum_{i=1}^{m} (f(x_i) - y_i)^2.$$

You must explain what the variables are and how they parametrize $f$, and how you ensure convexity of $f$.

*Hints.* One method to solve this problem is based on the Lagrange basis, $f_0, \ldots, f_K$, which are the piecewise linear functions that satisfy

$$f_j(a_i) = \delta_{ij}, \quad i, j = 0, \ldots, K.$$

Another method is based on defining $f(x) = \alpha_i x + \beta_i$, for $x \in (a_{i-1}, a_i]$. You then have to add conditions on the parameters $\alpha_i$ and $\beta_i$ to ensure that $f$ is continuous and convex.

Apply your method to the data in the file `pwl_fit_data.m`, which contains data with $x_j \in [0, 1]$. Find the best affine fit (which corresponds to $a = (0, 1)$), and the best piecewise-linear convex function fit for 1, 2, and 3 internal knot points, evenly spaced in $[0, 1]$. (For example, for 3 internal knot points we have $a_0 = 0,\ a_1 = 0.25,\ a_2 = 0.50,\ a_3 = 0.75,\ a_4 = 1$.) Give the least-squares fitting cost for each one. Plot the data and the piecewise-linear fits found. Express each function in the form

$$f(x) = \max_{i=1\ldots,K} (\alpha_i x + \beta_i).$$

(In this form the function is easily incorporated into an optimization problem.)

4. *Robust least-squares with interval coefficient matrix.* An *interval matrix* in $\mathbf{R}^{m \times n}$ is a matrix whose entries are intervals:

$$\mathcal{A} = \{A \in \mathbf{R}^{m \times n} \mid |A_{ij} - \bar{A}_{ij}| \leq R_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n\}.$$

The matrix $\bar{A} \in \mathbf{R}^{m \times n}$ is called the *nominal value* or *center value*, and $R \in \mathbf{R}^{m \times n}$, which is elementwise nonnegative, is called the *radius*.

The robust least-squares problem, with interval matrix, is

$$\text{minimize} \quad \sup_{A \in \mathcal{A}} \|Ax - b\|_2,$$

with optimization variable $x \in \mathbf{R}^n$. The problem data are $\mathcal{A}$ (*i.e.*, $\bar{A}$ and $R$) and $b \in \mathbf{R}^m$. The objective, as a function of $x$, is called the *worst-case residual norm.* The robust least-squares problem is evidently a convex optimization problem.

(a) Formulate the interval matrix robust least-squares problem as a standard optimization problem, *e.g.*, a QP, SOCP, or SDP. You can introduce new variables if needed. Your reformulation should have a number of variables and constraints that grows linearly with $m$ and $n$, and not exponentially.

(b) Consider the specific problem instance with $m = 4$, $n = 3$,

$$\mathcal{A} = \begin{bmatrix} 60 \pm 0.05 & 45 \pm 0.05 & -8 \pm 0.05 \\ 90 \pm 0.05 & 30 \pm 0.05 & -30 \pm 0.05 \\ 0 \pm 0.05 & -8 \pm 0.05 & -4 \pm 0.05 \\ 30 \pm 0.05 & 10 \pm 0.05 & -10 \pm 0.05 \end{bmatrix}, \qquad b = \begin{bmatrix} -6 \\ -3 \\ 18 \\ -9 \end{bmatrix}.$$

(The first part of each entry in $\mathcal{A}$ gives $\bar{A}_{ij}$; the second gives $R_{ij}$, which are all 0.05 here.) Find the solution $x_{\rm ls}$ of the nominal problem (*i.e.*, minimize $\|\bar{A}x - b\|_2$), and robust least-squares solution $x_{\rm rls}$. For each of these, find the nominal residual norm, and also the worst-case residual norm. Make sure the results make sense.

5. *Total variation image interpolation.* A grayscale image is represented as an $m \times n$ matrix of intensities $U^{\rm orig}$. You are given the values $U^{\rm orig}_{ij}$, for $(i,j) \in \mathcal{K}$, where $\mathcal{K} \subset \{1,\ldots,m\} \times \{1,\ldots,n\}$. Your job is to *interpolate* the image, by guessing the missing values. The reconstructed image will be represented by $U \in \mathbf{R}^{m\times n}$, where $U$ satisfies the interpolation conditions $U_{ij} = U^{\rm orig}_{ij}$ for $(i,j) \in \mathcal{K}$.

   The reconstruction is found by minimizing a roughness measure subject to the interpolation conditions. One common roughness measure is the $\ell_2$ variation (squared),

$$\sum_{i=2}^{m}\sum_{j=2}^{n}\left((U_{ij} - U_{i-1,j})^2 + (U_{ij} - U_{i,j-1})^2\right).$$

   Another method minimizes instead the *total variation*,

$$\sum_{i=2}^{m}\sum_{j=2}^{n}\left(|U_{ij} - U_{i-1,j}| + |U_{ij} - U_{i,j-1}|\right).$$

   Evidently both methods lead to convex optimization problems.

   Carry out $\ell_2$ and total variation interpolation on the problem instance with data given in `tv_img_interp.m`. This will define `m`, `n`, and matrices `Uorig` and `Known`. The matrix `Known` is $m \times n$, with $(i,j)$ entry one if $(i,j) \in \mathcal{K}$, and zero otherwise. The mfile also has skeleton plotting code. (We give you the entire original image so you can compare your reconstruction to the original; obviously your solution cannot access $U^{\rm orig}_{ij}$ for $(i,j) \notin \mathcal{K}$.)

6. *Relaxed and discrete A-optimal experiment design.* This problem concerns the $A$-optimal experiment design problem, described on page 387, with data generated as follows.

```
n = 5; % dimension of parameters to be estimated
p = 20; % number of available types of measurements
m = 30; % total number of measurements to be carried out
randn('state', 0);
V=randn(n,p); % columns are vi, the possible measurement vectors
```

   Solve the relaxed $A$-optimal experiment design problem,

$$\begin{array}{ll} \text{minimize} & (1/m)\,\mathbf{tr}\left(\sum_{i=1}^{p} \lambda_i v_i v_i^T\right)^{-1} \\ \text{subject to} & \mathbf{1}^T\lambda = 1, \quad \lambda \succeq 0, \end{array}$$

with variable $\lambda \in \mathbf{R}^p$. Find the optimal point $\lambda^\star$ and the associated optimal value of the relaxed problem. This optimal value is a lower bound on the optimal value of the discrete $A$-optimal experiment design problem,

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}\left(\sum_{i=1}^p m_i v_i v_i^T\right)^{-1} \\ \text{subject to} & m_1 + \cdots + m_p = m, \quad m_i \in \{0, \ldots, m\}, \quad i = 1, \ldots, p, \end{array}$$

with variables $m_1, \ldots, m_p$. To get a suboptimal point for this discrete problem, round the entries in $m\lambda^\star$ to obtain integers $\hat{m}_i$. If needed, adjust these by hand or some other method to ensure that they sum to $m$, and compute the objective value obtained. This is, of course, an upper bound on the optimal value of the discrete problem. Give the gap between this upper bound and the lower bound obtained from the relaxed problem. Note that the two objective values can be interpreted as mean-square estimation error $\mathbf{E}\,\|\hat{x} - x\|_2^2$.