SMOOTHED FIT TO GIVEN DATA (10 points possible)

Consider the problem

$$\text{minimize} \quad f(x) = \sum_{i=1}^{n} \psi(x_i - y_i) + \lambda \sum_{i=1}^{n-1}(x_{i+1} - x_i)^2$$

where $\lambda > 0$ is smoothing parameter, ψ is a convex penalty function, and $x \in \mathbf{R}^n$ is the variable. We can interpret x as a smoothed fit to the vector y. You can assume ψ is amooth.

What is the structure in the Hessian of f?

- ○ diagonal
- ○ diagonal + rank one
- ○ tridiagonal
- ○ tridiagonal + low rank
- ○ pentadiagonal
- ○ circulant

Show Answer   *You have used 0 of 2 submissions*

MAXIMUM LIKELIHOOD PREDICTION OF TEAM ABILITY (20 points possible)

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 $\mathbf{prob}(a_j - a_k + v > 0)$, 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.)

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.

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.) You may find the CVX function `log_normcdf` helpful for this problem.

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);
```

Finally, use the maximum likelihood estimate $\hat{a}$ 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`. What is the fraction of correctly predicted outcomes? (If half of the outcomes are correctly predicted, the answer is 0.5 and not 50.)

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

Show Answer   *You have used 0 of 2 submissions*

ALLOCATION OF INTERDICTION EFFORT (20 points possible)

A smuggler moves along a directed acyclic graph with m edges and n nodes, from a source node (which we take as node 1) to a destination node (which we take as node n), along some (directed) path. Each edge k has a detection failure probability p_k, which is the probability that the smuggler passes over that edge undetected. The detection events on the edges are independent, so the probability that the smuggler makes it to the destination node undetected is $\prod_{j \in \mathcal{P}} p_j$, where $\mathcal{P} \subset \{1, \ldots, m\}$ is (the set of edges on) the smuggler's path. We assume that the smuggler knows the detection failure probabilities and will take a path that maximizes the probability of making it to the destination node undetected. We let $P^{\max}$ denote this maximum probability (over paths). (Note that this is a function of the edge detection failure probabilities.)

The edge detection failure probability on an edge depends on how much interdiction resources are allocated to the edge. Here we will use a very simple model, with $x_j \in \mathbf{R}_+$ denoting the effort (say, yearly budget) allocated to edge j, with associated detection failure probability $p_j = e^{-a_j x_j}$, where $a_j \in \mathbf{R}_{++}$ are given. The constraints on x are a maximum for each edge, $x \preceq x^{\max}$, and a total budget constraint, $\mathbf{1}^T x \leq B$.

Pose the problem of choosing the interdiction effort vector $x \in \mathbf{R}^m$ that minimizes P^{max}, subject to the constraints, as a convex optimization problem. Your formulation should not enumerate all possible paths (in the objective or constraints). *Hint.* For each node i, let P_i denote the maximum of $\prod_{k \in \mathcal{P}} p_k$ over all paths $\mathcal{P}$ from the source node 1 to node i (so $P^{\mathrm{max}} = P_n$).

Carry out your method on the problem instance given in `interdict_alloc_data.m`. The data file contains the data a, x^{max}, B, and the graph incidence matrix $A \in \mathbf{R}^{n \times m}$, where

$$
A_{ij} = \begin{cases} -1 & \text{if edge } j \text{ leaves node } i \\ +1 & \text{if edge } j \text{ enters node } i \\ 0 & \text{otherwise.} \end{cases}
$$

Give $P^{\mathrm{max}\star}$, the optimal value of P^{max}, and compare it to the value of P^{max} obtained with uniform allocation of resources, i.e., with $x = (B/m)\mathbf{1}$.

Hint. Given a vector $z \in \mathbf{R}^n$, $A^T z$ is the vector of edge differences: $(A^T z)_j = z_k - z_l$ if edge j goes from node l to node k.
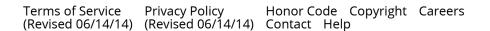
What is the value of $P^{\mathrm{max}\star}$? (Write 0.1 for 10%.)

What is the value of P^{max} obtained with uniform allocation of resources? (Write 0.1 for 10%.)

Show Answer   *You have used 0 of 2 submissions*