EE364b Spring 2020 Homework 5

Due Friday 5/7 at 11:59pm via Gradescope

5.1 *(4 points) Interior Point Method for Lasso.* In this problem, you will implement a basic barrier method for solving the Lasso problem

$$\min_x \frac{1}{2}\|Ax - y\|_2^2 + \beta\|x\|_1$$

(a) Derive the dual problem. Show that the dual is a Linearly Constrained Quadratic Program (LCQP), i.e., in the following form

$$\begin{aligned} &\min_\lambda \frac{1}{2}\lambda^T Q\lambda + q^T\lambda \\ &\text{s.t. } B\lambda \le c \end{aligned}$$

where Q, B are matrices and q, c are vectors of appropriate size.

(b) Generate random standard Gaussian matrix $A \in \mathbb{R}^{10\times 100}$ and vector $y \in \mathbb{R}^{10}$ with independent entries. Set $\beta = 1$. Implement Newton's Method with Armijo line-search to minimize the log-barrier penalized centering problem

$$\min_\lambda t\Big(\frac{1}{2}\lambda^T Q\lambda + q^T\lambda\Big) + \sum_{i=1}^n -\log(c_i - B_i^T\lambda)\,.$$

where B_i is the i'th row of the matrix B and c_i is the i'th element of the vector c. Set $t = 1000$. Plot the objective value of the centering problem vs iteration number.

5.2 *(4 points) Maximum volume ellipsoid vs Chebyshev center method.* Consider the convex set

$$\mathcal{C} = \{x \mid Ax \preceq b\}\,,$$

where $A \in \mathbf{R}^{n\times d}$ and $b \in \mathbf{R}^d$. The data files `Amatrix` and `bvector` are available on Canvas.

(a) (2 points) Find the center of the maximum volume ellipsoid in $\mathcal{C}$ and the center of the largest Euclidean ball in $\mathcal{C}$. You may use CVX/CVXPY. *Hint: See 364a slides* for calculating the maximum volume ellipsoid.

(b) (2 points) Denote the two centers (vectors in $\mathbf{R}^d$) in part (a) by $x_{\text{ellipsoid}}$ and x_{ball} respectively. Let $g \in \mathbf{R}^d$ be the all-ones vector. We will consider the cuts $g^T(x - x_{\text{ball}}) \geq 0$ and $g^T(x - x_{\text{ellipsoid}}) \geq 0$. Estimate the volume ratios

$$R_{\text{ellipsoid}} := \frac{\mathbf{vol}(\{g^T(x - x_{\text{ellipsoid}}) \geq 0)\} \cap \mathcal{C})}{\mathbf{vol}(\mathcal{C})},$$

and

$$R_{\text{ball}} := \frac{\mathbf{vol}(\{g^T(x - x_{\text{ball}}) \geq 0)\} \cap \mathcal{C})}{\mathbf{vol}(\mathcal{C})},$$

by generating $M = 10^6$ i.i.d. uniformly distributed random vectors in $[-0.5, +0.5]^d$ (i.e., $x =$ `rand(d,1)-0.5` for M trials). *Hint:* Let $M_{\mathcal{C}}$ be number of random vectors that satisfy $Ax \preceq b$. Let $M_{\text{ellipsoid}}$ be the number of random vectors that satisfy $Ax \preceq b$ and $g^T(x - x_{\text{ellipsoid}}) \geq 0$. Similarly, let M_{ball} be the number of random vectors that satisfy $Ax \preceq b$ and $g^T(x - x_{\text{ball}}) \geq 0$. The volume ratios can be estimated by

$$R_{\text{ellipsoid}} \approx \frac{M_{\text{ellipsoid}}}{M_{\mathcal{C}}},$$

and

$$R_{\text{ball}} \approx \frac{M_{\text{ball}}}{M_{\mathcal{C}}}.$$

5.3 *(5 points) Kelley's cutting-plane algorithm.* We consider the problem of minimizing a convex function $f : \mathbf{R}^n \to \mathbf{R}$ over some convex set C, assuming we can evaluate $f(x)$ and find a subgradient $g \in \partial f(x)$ for any x. Suppose we have evaluated the function and a subgradient at $x^{(1)}, \ldots, x^{(k)}$. We can form the piecewise-linear approximation

$$\hat{f}^{(k)}(x) = \max_{i=1,\ldots,k} \left(f(x^{(i)}) + g^{(i)T}(x - x^{(i)}) \right),$$

which satisfies $\hat{f}^{(k)}(x) \leq f(x)$ for all x. It follows that

$$L^{(k)} = \inf_{x \in C} \hat{f}^{(k)}(x) \leq p^\star,$$

where $p^\star = \inf_{x \in C} f(x)$. Since $\hat{f}^{(k+1)}(x) \geq \hat{f}^{(k)}(x)$ for all x, we have $L^{(k+1)} \geq L^{(k)}$.

In Kelley's cutting-plane algorithm, we set $x^{(k+1)}$ to be any point that minimizes $\hat{f}^{(k)}$ over $x \in C$. The algorithm can be terminated when $U^{(k)} - L^{(k)} \leq \epsilon$, where $U^{(k)} = \min_{i=1,\ldots,k} f(x^{(i)})$.

(a) (3 points) Use Kelley's cutting-plane algorithm to minimize the piecewise-linear function

$$f(x) = \max_{i=1,\ldots,m} (a_i^T x + b_i)$$

that we have used for other numerical examples, with C the unit cube, *i.e.*, $C = \{x \mid \|x\|_\infty \leq 1\}$. Generate the same data we used before using

```
n = 20; % number of variables
m = 100; % number of terms
randn('state',1);
A = randn(m,n);
b = randn(m,1);
```

You can start with $x^{(1)} = 0$ and run the algorithm for 40 iterations. Plot $f(x^{(k)})$, $U^{(k)}$, $L^{(k)}$ and the constant $p^\star$ (on the same plot) versus k.

(b) (2 points) Repeat for $f(x) = \|x - c\|_2$, where c is chosen from a uniform distribution over the unit cube C. (The solution to this problem is, of course, $x^\star = c$.)

5.4 *(Extra credit, 6 points) Minimum volume ellipsoid covering a half-ellipsoid.* In this problem we derive the update formulas used in the ellipsoid method, *i.e.*, we will determine the minimum volume ellipsoid that contains the intersection of the ellipsoid

$$\mathcal{E} = \{x \in \mathbf{R}^n \mid (x - x_c)^T P^{-1}(x - x_c) \le 1\}$$

and the halfspace

$$\mathcal{H} = \{x \mid g^T(x - x_c) \le 0\}.$$

We'll assume that $n > 1$, since for $n = 1$ the problem is easy.

(a) (2 points) We first consider a special case: $\mathcal{E}$ is the unit ball centered at the origin ($P = I$, $x_c = 0$), and $g = -e_1$ (e_1 is the first unit vector), so $\mathcal{E} \cap \mathcal{H} = \{x \mid x^T x \le 1,\ x_1 \ge 0\}$.

Let

$$\tilde{\mathcal{E}} = \{x \mid (x - \tilde{x}_c)^T \tilde{P}^{-1}(x - \tilde{x}_c) \le 1\}$$

denote the minimum volume ellipsoid containing $\mathcal{E} \cap \mathcal{H}$. Since $\mathcal{E} \cap \mathcal{H}$ is symmetric about the line through first unit vector e_1, it is clear (and not too hard to show) that $\tilde{\mathcal{E}}$ will have the same symmetry. This means that the matrix $\tilde{P}$ is diagonal, of the form $\tilde{P} = \mathbf{diag}(\alpha, \beta, \beta, \ldots, \beta)$, and that $\tilde{x}_c = \gamma e_1$ (where $\alpha, \beta > 0$ and $\gamma \ge 0$).

So now we have only three variables to determine: α, β, and γ. Express the volume of $\tilde{\mathcal{E}}$ in terms of these variables, and also the constraint that $\tilde{\mathcal{E}} \supseteq \mathcal{E} \cap \mathcal{H}$. Then solve the optimization problem directly, to show that

$$\alpha = \frac{n^2}{(n+1)^2}, \quad \beta = \frac{n^2}{n^2 - 1}, \quad \gamma = \frac{1}{n+1}$$

(which agrees with the formulas we gave, for this special case).

Hint. To express $\mathcal{E} \cap \mathcal{H} \subseteq \tilde{\mathcal{E}}$ in terms of the variables, it is necessary and sufficient for the conditions on α, β, and γ to hold on the boundary of $\mathcal{E} \cap \mathcal{H}$, *i.e.*, at the points

$$x_1 = 0, \quad x_2^2 + \cdots + x_n^2 \le 1,$$

or the points

$$x_1 \ge 0, \quad x_1^2 + x_2^2 + \cdots + x_n^2 = 1.$$

(b) (2 points) Now consider the general case, stated at the beginning of this problem. Show how to reduce the general case to the special case solved in part (a).

Hint. Find an affine transformation that maps the original ellipsoid to the unit ball, and g to $-e_1$. Explain why minimizing the volume in these transformed coordinates also minimizes the volume in the original coordinates.

(c) (2 points) Finally, show that the volume of the ellipse $\tilde{\mathcal{E}}$ satisfies $\mathbf{vol}(\tilde{\mathcal{E}}) \leq e^{-\frac{1}{2n}} \mathbf{vol}(\mathcal{E})$.

Hint. Compute the volume of the ellipse $\mathcal{E}$ as a function of the eigenvalues of P, then use the results of parts (a) and (b) to argue that the volume computation can be reduced to the special case in part (a).