EE364b Spring 2020 Homework 5
Due Friday 5/15 at 11:59pm via Gradescope

4.1 (4 points) Maximum volume ellipsoid vs Chebyshev center method. Consider the convex set
\[ C = \{ x \mid Ax \preceq b \} , \]
where \( A \in \mathbb{R}^{n \times d} \) and \( b \in \mathbb{R}^d \). The data files \( \text{Amatrix} \) and \( \text{bvector} \) are available on Canvas.

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

(b) (2 points) Denote the two centers (vectors in \( \mathbb{R}^d \)) in part (a) by \( x_{\text{ellipsoid}} \) and \( x_{\text{ball}} \) respectively. Let \( g \in \mathbb{R}^d \) 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
\[
R_{\text{ellipsoid}} := \frac{\text{vol}(\{ (g^T(x - x_{\text{ellipsoid}}) \geq 0) \cap C \})}{\text{vol}(C)},
\]
and
\[
R_{\text{ball}} := \frac{\text{vol}(\{ (g^T(x - x_{\text{ball}}) \geq 0) \cap C \})}{\text{vol}(C)},
\]
by generating \( M = 10^6 \) i.i.d. uniformly distributed random vectors in \([ -0.5, +0.5]^d \) (i.e., \( x = \text{rand}(d,1) - 0.5 \) for \( M \) trials). \textit{Hint: Let} \( M_C \) \text{ be number of random vectors that satisfy } Ax \preceq b. \text{ Let } M_{\text{ellipsoid}} \text{ be the number of random vectors that satisfy } Ax \preceq b \text{ and } g^T(x - x_{\text{ellipsoid}}) \geq 0. \text{ Similarly, let } M_{\text{ball}} \text{ be the number of random vectors that satisfy } Ax \preceq b \text{ and } g^T(x - x_{\text{ball}}) \geq 0. \text{ The volume ratios can be estimated by}
\[
R_{\text{ellipsoid}} \approx \frac{M_{\text{ellipsoid}}}{M_C},
\]
and
\[
R_{\text{ball}} \approx \frac{M_{\text{ball}}}{M_C}.
\]
4.2 (5 points) Kelley’s cutting-plane algorithm. We consider the problem of minimizing a convex function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) 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

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

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

\[
L^{(k)} = \inf_{x \in C} \hat{f}^{(k)}(x) \leq p^*,
\]

where \( p^* = \inf_{x \in C} f(x) \). 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

\[
f(x) = \max_{i=1, \ldots, m} \left( a_i^T x + b_i \right)
\]

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

\[
n = 20; \quad \% \text{ number of variables}
\]
\[
m = 100; \quad \% \text{ number of terms}
\]
\[
\text{randn('state',1)};
\]
\[
A = \text{randn}(m,n);
\]
\[
b = \text{randn}(m,1);
\]

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 \( p^* \) (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, \( x^* = c \).)

4.3 (5 points) Ellipsoid method for an SDP. We consider the SDP

\[
\begin{align*}
\text{maximize} & \quad 1^T x
\end{align*}
\]

\[
\text{subject to} \quad x \succeq 0, \quad \Sigma - \text{diag}(x) \succeq 0,
\]

with variable \( x \in \mathbb{R}^n \) and data \( \Sigma \in \mathbb{S}^n_+ \). The first inequality is a vector (component-wise) inequality, and the second inequality is a matrix inequality.
(a) (3 points) Explain how to use the ellipsoid method to solve this problem. Describe your choice of initial ellipsoid and how you determine a subgradient for the objective (expressed as $-1^T x$, which is to be minimized) or constraint functions (expressed as $\max_i(-x_i) \leq 0$ and $\lambda_{\text{max}}(\text{diag}(x) - \Sigma) \leq 0$). You can describe a basic ellipsoid method; you do not need to use a deep-cup method or work in the epigraph.

(b) (2 points) Try out your ellipsoid method on some randomly generated data with $n \leq 20$. Use a stopping criterion that guarantees 1% accuracy. Compare the result of the solution found using CVX. Plot the upper and lower bounds from the ellipsoid method, versus iteration number.

4.4 (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

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

and the halfspace

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

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 $\mathcal{E} \cap \mathcal{H} = \{ x \mid x^T x \leq 1, \ x_1 \geq 0 \}$.

Let

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

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 $\tilde{P} = \text{diag}(\alpha, \beta, \beta, \ldots, \beta)$, and that $\tilde{x}_c = \gamma e_1$ (where $\alpha, \beta > 0$ and $\gamma \geq 0$).

So now we have only three variables to determine: $\alpha$, $\beta$, and $\gamma$. 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 $\alpha$, $\beta$, and $\gamma$ to hold on the boundary of $\mathcal{E} \cap \mathcal{H}$, i.e., at the points

$$x_1 = 0, \quad x_2^2 + \cdots + x_n^2 \leq 1,$$
or the points
\[ x_1 \geq 0, \quad x_1^2 + x_2^2 + \cdots + x_n^2 = 1. \]

(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 \( \tilde{E} \) satisfies \( \text{vol}(\tilde{E}) \leq e^{-\frac{1}{n}} \text{vol}(E) \).

**Hint.** Compute the volume of the ellipse \( E \) 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).