ISyE6669 Homework 2
September 11, 2018
On-Campus Deadline: September 24, 2018
DL Deadline: October 1, 2018

1 Convex Sets

1. Prove that if sets S_1 and S_2 are convex sets in $\mathbb{R}^n$, then their intersection $S_1 \cap S_2$ is a convex set, using the definition of convex sets.

   Hint: This should be an easy argument following from the definition. Also, we always take an empty set as a convex set.

   Remark: This exercise tells us an important property that the intersection of two convex sets is a convex set, in short we say, intersection *preserves* convexity.

2. Extend the above result to argue that $S_1 \cap S_2 \cap ... \cap S_m$ is convex if all $S_1, \dots, S_m$ are convex sets in $\mathbb{R}^n$ for any $m \geq 1$.

3. Show that a halfspace $H^+ = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} \geq \boldsymbol{b}\}$ is a convex set using the definition of a convex set.

4. Show that a hyperplane $H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} = \boldsymbol{b}\}$ is a convex set by using the conclusion of questions 1 and 3.

5. Argue that a polyhedron is always a convex set by using the results from questions 2 and 3.

6. Show that the convex hull S of a finite number of points $\boldsymbol{x}^1, \dots, \boldsymbol{x}^m$ in $\mathbb{R}^n$ is a convex set. Hint: Use the definition of a convex set.

7. Give an example to show that a nonempty convex set may not be the convex hull of a finite number of points.

8. Draw the following sets by hands using ruler and pencil (your circles do not have to be perfect). State if each set is a convex set, and list all the extreme points of each set. If a set does not have an extreme point, just say so. Hint: The definition of an extreme point applies to nonconvex sets.

   (a) $T_1 = \{\boldsymbol{x} \in \mathbb{R}^2 : x_1 + x_2 \leq 1\}$.

   (b) $T_2 = \text{conv}\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$.

   (c) $T_3 = \{\boldsymbol{x} \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1\}$.

   (d) $T_4 = \{\boldsymbol{x} \in \mathbb{R}^3 : |x_1| + |x_2| + |x_3| \leq 1\}$.

   (e) $T_5 = \{\boldsymbol{x} \in \mathbb{R}^2 : 1 \leq x_1^2 + x_2^2 \leq 4\}$.

   (f) $T_6 = \{0, 1, 2, 3, 4\} \subset \mathbb{R}$.

2 Convex Functions

1. Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ that is both convex and concave.

2. Let $f_1 : \mathbb{R}^n \to \mathbb{R}$ and $f_2 : \mathbb{R}^n \to \mathbb{R}$ be two convex functions. Prove that the sum $f = f_1 + f_2$ is always a convex function. What about $af_1 + bf_2$ for any nonnegative constants a, b? Hint: This should be a straightforward argument from the Jensen's inequality, and this shows linear combination of convex functions with nonnegative weights is again a convex function.

3. Let $f_1 = |x|$ and $f_2 = |x+1|$. Notice that both f_1 and f_2 are convex functions. Let $f = f_1 - f_2$. Draw $f_1(x)$, $f_2(x)$, and $f(x)$ for $x \in [-2, 2]$ on the same plot *by hand*, i.e. using ruler and pencil on a piece of paper by hand. Is f a convex function on $[-2, 2]$? Hint: Your answer should show that the *difference* of two convex functions may not be a convex function any more. Contrast this with the previous question.

4. Let $f_0 = x^2, f_1 = (x-1)^2, ..., f_k = (x-k)^2$ for all $x \in \mathbb{R}$. Let $g_k(x) = \max\{f_1(x), ..., f_k(x)\}$ for all $x \in \mathbb{R}$. Draw f_0, f_1, f_2, and $g_1(x)$ and $g_2(x)$ on two separate plots. Are they convex sets?

5. Let $f_1(x), ..., f_k(x)$ be convex functions on $\mathbb{R}$. Prove $g(x) = \max\{f_1(x), \ldots, f_k(x)\}$ is a convex function. Hint: Use the Jensen's inequality and a similar argument that you used in Homework 1 for proving a similar result. This exercise shows that the pointwise max operation preserves convexity of functions.

6. Define a set $S = \{(x, y) \in \mathbb{R}^2 : y = \sin(x)\}$, i.e. S is the graph of the sine function on the entire real axis. Draw S. Is S a convex set in $\mathbb{R}^2$? Find the convex hull of S by drawing it. Write down the inequalities that define the convex hull of S. Hint: This exercise requires you to find the convex hull of an infinite number of points. Your intuition of how to form the convex hull of a finite number of points should suffice to guide you.

3 LP geometry and the simplex method

1. Consider the following linear program:

$$
\begin{array}{lrcrcl}
\min & -2x_1 & - & 3x_2 & & \\
\text{s.t.} & x_1 & + & x_2 & \leq & 4 \\
 & -x_1 & + & x_2 & \leq & 2 \\
 & x_1 & - & x_2 & \leq & 2 \\
 & \multicolumn{5}{l}{x_1 \geq 0, \quad x_2 \geq 0.}
\end{array}
$$

   Draw the feasible region of this linear program.

2. To solve this problem using the simplex method, first transform it into a standard form LP. Denote $\boldsymbol{x} = [x_1, x_2, x_3, x_4, x_5]^\top$ as the vector of variables, and use the standard form notation:

$$
\begin{array}{ll}
\min & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} \geq \boldsymbol{0},
\end{array}
$$

   specify $\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}$ for the above problem.

3. Now we want to solve the above standard-form linear program by the simplex method. If in an iteration of the Simplex method, there is any ambiguity about which nonbasic variable to enter the basis or which basic variable to exit the basis, use the Bland's rule.

   (a) For each iteration of the simplex method, write down the following information in the format given below:

   - Iteration $k = \underline{\text{numerical value, e.g. 1, 2, ...}}$;
   - Basis $\boldsymbol{B} = [\boldsymbol{A}_i, \boldsymbol{A}_j, \boldsymbol{A}_k] = \underline{\text{numerical value}}$ (i.e. you need to specify i, j, k as well as the numerical values of the columns);
   - Basis inverse $\boldsymbol{B}^{-1} = \underline{\text{numerical value}}$;
   - Basic variable $\boldsymbol{x}_B = [x_i, x_j, x_k] = \underline{\text{numerical value}}$ (you need to specify i, j, k and numerical values of x_i, x_j, x_k);
   - Nonbasic variable $\boldsymbol{x}_N = [x_p, x_q] = \underline{\text{numerical value}}$ (you need to specify p, q and numerical values of x_p, x_q);
   - Reduced cost for each nonbasic variable $\bar{c}_p = c_p - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1} \boldsymbol{A}_? = \underline{\text{numerical values}}$; Same for $\bar{c}_q$; (you need to the index "?" for $\boldsymbol{A}_?$);
   - Is the current solution optimal? If not, which nonbasic variable to enter the basis?
   - Direction $\boldsymbol{d}_B = -\boldsymbol{B}^{-1}\boldsymbol{A}_? = \underline{\text{numerical value}}$. Does the simplex method terminate with an unbounded optimum?
   - Min-ratio test $\theta^* = \min_{i:d_{B(i)}<0}\{x_{B(i)}/(-d_{B(i)})\} = \min\{\underline{\text{numerical values of the ratios}}\} = \underline{\text{numerical value of } \theta^*}$.
   - Which basic variable to exit the basis?

   Start at iteration $k = 1$ with the basis $\boldsymbol{B}^1 = [\boldsymbol{A}_3, \boldsymbol{A}_4, \boldsymbol{A}_5]$. Solve the above linear program with simplex method and write down all the information required above for each iteration. Also indicate the basic feasible solution at each step on the picture in (x_1, x_2). What is the optimal solution of the above LP in $(x_1, \ldots, x_5)$? What is the optimal cost?

From this exercise, you can see how the simplex method works and geometrically what each step is doing.

4 Modeling exercise: least squares and robust regressions

Given a set of training data $\{\boldsymbol{x}_i, y_i\}_{i=1,\ldots,N}$, where $\boldsymbol{x}_i$ is an n-dimensional feature vector and y_i is a label of value either 0 or 1. Think about each $\boldsymbol{x}_i$ represents a vector of lab test data of a patient i and y_i labels if this person has a certain disease. We want to build a linear classifier, i.e. a linear function $f(\boldsymbol{x}) = \beta_0 + \sum_{j=1}^n \beta_j x_j$, so that for a given feature vector $\boldsymbol{x}$, if $f(\boldsymbol{x}) \geq 0.5$, then $\boldsymbol{x}$ is classified as $y = 1$, otherwise classified as $y = 0$.

There are many ways to build this linear classifier. Most of them try to find the coefficients of the linear function by minimizing certain distance metric over the training data. A very popular method is called the absolute deviation regression (ADR). ADR is also called robust regression. The optimization model of ADR is described below.

$$\text{(ADR)} \quad \min_{\beta_0,\ldots,\beta_n} \sum_{i=1}^{N} \left| y_i - \beta_0 - \sum_{j=1}^{n} \beta_j x_{ij} \right|,$$

where x_{ij} is the jth component of vector $\boldsymbol{x}_i$ and both x_{ij}'s and y_i's are the given data, while β_i's are the decision variables. Note that the (ADR) model in the above form is a nonlinear optimization problem, which cannot be directly solved by Xpress.

Answer the following questions.

1. Is the objective function of (ADR) a convex function in $\beta_0, \ldots, \beta_n$? (The answer should be yes.) But explain why, using the conclusion of question 2.2.

2. Write down a linear programming reformulation of (ADR).

3. Code your LP reformulation of (ADR) in Xpress, using the data file and the partial mos file provided. Pay attention to the format of the data file and how it is handled in the mos file. Submit your mos code in t-square. And write down your solution in the hardcopy of your homework.

4. Use the MATLAB code to plot the hyperplane obtained from (ADR). Submit a hard copy of the plot.