ISyE6669 Deterministic Optimization Homework 2

Solutions

October 18, 2018

1 (25 pt) Convex Sets

1. (3 pt) For $\forall x_1, x_2 \in S_1 \cap S_2$ and $\forall x$ lies on the line segment between x_1, x_2, we have $x \in S_1$ because $x_1, x_2 \in S_1$ and S_1 is convex, $x \in S_2$ because $x_1, x_2 \in S_2$ and S_2 is convex. Thus, $x \in S_1 \cap S_2$, and we conclude that $S_1 \cap S_2$ is still a convex set.

2. (3 pt) Let $A_2 = S_1 \cap S_2$, and A_2 is convex by the results in1.1. Let $A_3 = S_1 \cap S_2 \cap S_3 = A_2 \cap S_3$, which is also convex following directly from the results in 1.1. By induction, assume $A_{m-1} = S_1 \cap \ldots \cap S_{m-1}$ is convex, then $A_m = A_{m-1} \cap S_m$ is still convex. Therefore, we conclude that $S_1 \cap \ldots \cap S_m$ is convex.

3. (3 pt) For $\forall x_1, x_2 \in H^+$, let x lies in the line segment between x_1, x_2, then x can be written as $x = \theta x_1 + (1-\theta)x_2$ $(0 \leq \theta \leq 1)$. Since $a'x = a'[\theta x_1 + (1-\theta)x_2] \leq \theta b + (1-\theta)b = b$, $x \in H^+$. Therefore, H^+ is a convext set.

4. (3 pt) Let $H^+ = \{x \in \mathbb{R}^n, a/x \geq b\}$, $H^- = \{x \in \mathbb{R}^n, a/x \leq b\}$. By 1.3, we know that H^+ is a convex set. Similarly, H^- is also a convex set. Let $H = H^+ \cap H^- = \{x \in \mathbb{R}^n, a/x = b\}$. By 1.1, H is also a convex set.

5. (3 pt)

   A polyhedron is an intersection of finite number of half spaces. We proved that all half space are convex in (3) and we proved that intersection of convex sets is also a convex set in (2). Hence, a polyhedron is also convex.

6. (3 pt) For $\forall x_1, x_2 \in \text{conv}\{x^1, \ldots, x^m\}$, let x lies in the line segment between x_1, x_2, i.e., $x = \theta x_1 + (1-\theta)x_2$, $0 \leq \theta \leq 1$.

$$\begin{aligned} x &= \theta x_1 + (1-\theta)x_2 \\ &= \theta(\sum_{i=1}^m a^i x^i) + (1-\theta)(\sum_{i=1}^m b^i x^i) \qquad \sum_{i=1}^m a^i = 1, \sum_{i=1}^m b^i = 1 \\ &= [\theta a^1 + (1-\theta)b^1]x^1 + \ldots + [\theta a^m + (1-\theta)b^m]x^m \end{aligned}$$

   Besides, $[\theta a^1 + (1-\theta)b^1] + \ldots + [\theta a^m + (1-\theta)b^m] = \theta \sum_{i=1}^m a^i + (1-\theta)\sum_{i=1}^m b^i = 1$. Therefore, x can be expressed by the convex combination of these m points. $x \in \text{conv}\{x^1, \ldots, x^m\}$.

7. (3 pt) example: $S = \{x = \{x_1, x_2\} : x_1^2 + x_2^2 \leq 1\}$

8. (4 pt)

   (a) convex, no extreme points

   (b) convex, extreme point: (-1,1), (-2,0), (-1,-1), (1,-1), (2,0), (1,1)

   (c) convex, extreme point: (0,1), (-1,0), (1,0), (0,-1)

   (d) convex, extreme point: (0,0,1), (1,0,0), (0,0,-1), (-1,0,0), (0,1,0), (0,-1,0)

   (e) nonconvex, extreme point: $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 4\}$

   (f) nonconvex, 0, 4

2 (25 pt) Convex Functions

1. (4 pt) Let $f(x) = x$, $x \in \mathbb{R}$. Then, f is both convex and concave on $\mathbb{R}$ because for any $x, y \in \mathbb{R}$ and $\lambda \in [0, 1]$,

$$f(\lambda x + (1-\lambda)y) = \lambda x + (1-\lambda)y = \lambda f(x) + (1-\lambda)f(y).$$

   (The equality satisfies the definitions of both convexity and concavity.)

2. (5 pt) (1) For $\forall x, y \in \mathbb{R}^n$,

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= f_1(\lambda x + (1-\lambda)y) + f_2(\lambda x + (1-\lambda)y) \\ &\leq \lambda f_1(x) + (1-\lambda)f_1(y) + \lambda f_2(x) + (1-\lambda)f_2(y) \\ &= \lambda[f_1(x) + f_2(x)] + (1-\lambda)[f_1(y) + f_2(y)] \\ &= \lambda f(x) + (1-\lambda)f(y) \end{aligned}$$

   Therefore, $f = f_1 + f_2$ is still convex.

   (2)For $\forall x, y \in \mathbb{R}^n$,

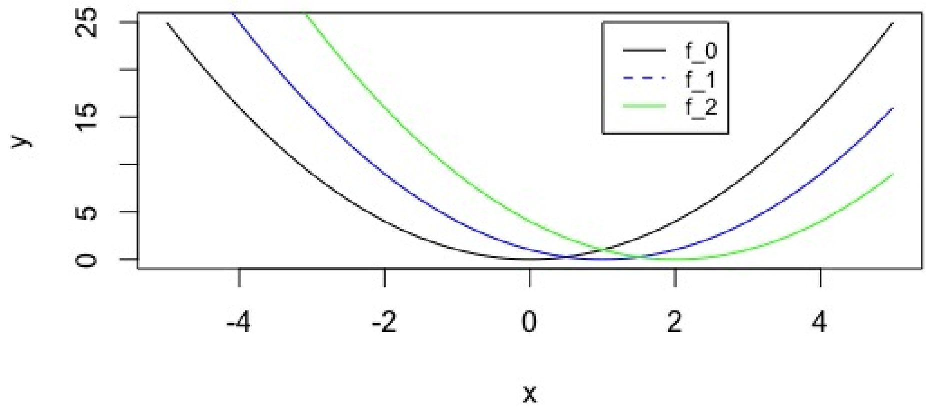
$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= af_1(\lambda x + (1-\lambda)y) + bf_2(\lambda x + (1-\lambda)y) \\ &\leq af_1(\lambda x) + af_1((1-\lambda)y) + bf_2(\lambda x) + bf_2((1-\lambda)y) \\ &= f(\lambda x) + f((1-\lambda)y) \end{aligned}$$

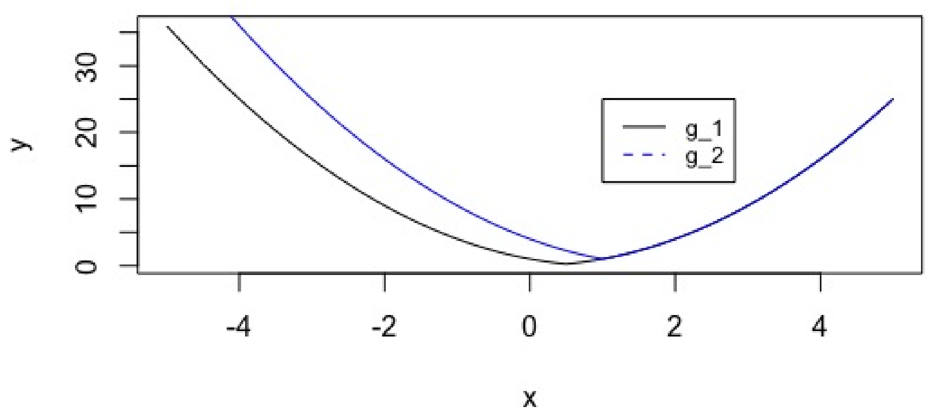
   Therefore, $f = af_1 + bf_2$ is still convex.

3. (4 pt)

$$f = \begin{cases} 1, & x < -1 \\ -2x - 1, & -1 \leq x \leq 0 \\ -1, & x > 0 \end{cases}$$

   f is no longer convex: Let $x_1 = -2, x_2 = 0, \lambda = \frac{1}{2}$, then $f(\lambda x + (1-\lambda)y) = f(-1) = 1$, but $\lambda f(x) + (1-\lambda)f(y) = \frac{1}{2} - \frac{1}{2} = 0$. $f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y)$. In the previous question, a, b are nonnegative constants, which ensures that the inequality holds when both sides times a nonnegative constant.





4. (4 pt) Yes, they are convex.

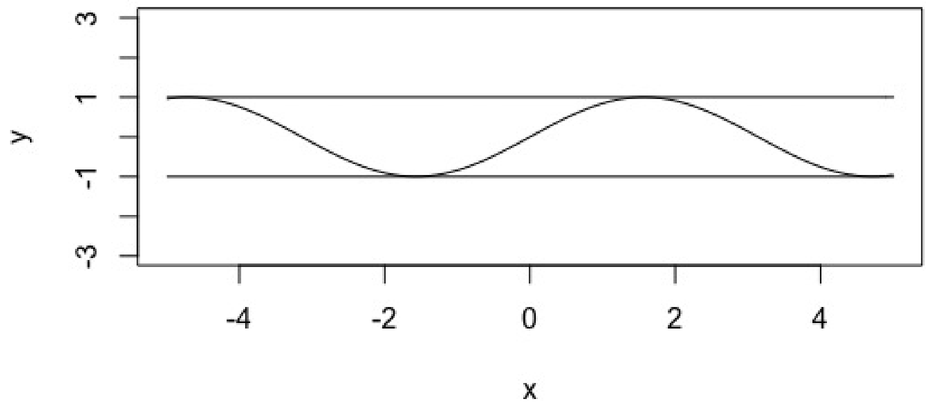
5. (4 pt) For $\forall x, y \in \mathbb{R}$,

$$\begin{aligned} g(\lambda x + (1-\lambda)y) &= \max\{f_1(\lambda x + (1-\lambda)y), \ldots, f_k(\lambda x + (1-\lambda)y)\} \\ &\leq \max\{\lambda f_1(x) + (1-\lambda) f_1(y), \ldots, \lambda f_k(x) + (1-\lambda) f_k(y)\} \\ &\leq \lambda \max\{f_1(x), \ldots, f_k(x)\} + (1-\lambda)\max\{f_1(y), \ldots, f_k(y)\} \\ &= \lambda g(x) + (1-\lambda) g(y) \end{aligned}$$

   $g(x)$ is convex.

6. (4 pt) The convex hull of S is $\text{conv}(S) = \{(x, y) \in \mathbb{R}^2 : -1 \leq y \leq 1\}$.

3 (25 pt) LP geometry and the simplex method



1. (5 pt)

2. (5 pt) The standard form can be written as

$$\begin{array}{lrcrcrcl} \min & -2x_1 & - & 3x_2 & & & & \\ \text{s.t.} & x_1 & + & x_2 & + & x_3 & = & 4 \\ & -x_1 & + & x_2 & + & x_4 & = & 2 \\ & x_1 & - & x_2 & + & x_5 & = & 2 \end{array}$$

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0, \quad x_5 \geq 0.$$

$$\mathbf{c} = \begin{bmatrix} -2 \\ -3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}$$

3. (15 pt) We solve the linear program with simplex method. To simplify the notation, we only specify the iteration index $k = 1, 2, \ldots$ at the beginning of each iteration step.

   When $k = 1$,

   - $\mathbf{B} = [\mathbf{A}_3, \mathbf{A}_4, \mathbf{A}_5] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
   - $\mathbf{x}_B = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}$, $\mathbf{x}_N = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

- $\bar{c}_1 = c_1 - \mathbf{c}_B^\mathsf{T}\mathbf{B}^{-1}\mathbf{A}_1 = -2,\ \bar{c}_2 = c_2 - \mathbf{c}_B^\mathsf{T}\mathbf{B}^{-1}\mathbf{A}_2 = -3.$
- $\bar{c}_1, \bar{c}_2 < 0 \implies$ the current solution is not optimal; the nonbasic variable x_1 is to enter the basis.
- $\mathbf{d}_B = -\mathbf{B}^{-1}\mathbf{A}_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$. The simplex method does not terminate with an unbounded optimum.
- $\theta^* = \min_{i:d_B(i)<0}\{x_{B(i)}/(-d_{B(i)})\} = 2.$
- The basis variable x_5 is to exit the basis.

When $k = 2$,

- $\mathbf{B} = [\mathbf{A}_3, \mathbf{A}_4, \mathbf{A}_1] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.
- $\mathbf{x}_B = \begin{bmatrix} x_3 \\ x_4 \\ x_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$, $\mathbf{x}_N = \begin{bmatrix} x_2 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
- $\bar{c}_2 = c_2 - \mathbf{c}_B^\mathsf{T}\mathbf{B}^{-1}\mathbf{A}_2 = -5,\ \bar{c}_5 = c_5 - \mathbf{c}_B^\mathsf{T}\mathbf{B}^{-1}\mathbf{A}_5 = 2.$
- $\bar{c}_2 < 0 \implies$ the current solution is not optimal; the nonbasic variable x_2 is to enter the basis.
- $\mathbf{d}_B = -\mathbf{B}^{-1}\mathbf{A}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$. The simplex method does not terminate with an unbounded optimum.
- $\theta^* = \min_{i:d_B(i)<0}\{x_{B(i)}/(-d_{B(i)})\} = 1.$
- The basis variable x_3 is to exit the basis.

When $k = 3$,

- $\mathbf{B} = [\mathbf{A}_2, \mathbf{A}_4, \mathbf{A}_1] = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$, $\mathbf{B}^{-1} = \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 1 & 1 \\ 1/2 & 0 & 1/2 \end{bmatrix}$.
- $\mathbf{x}_B = \begin{bmatrix} x_2 \\ x_4 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{x}_N = \begin{bmatrix} x_3 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
- $\bar{c}_3 = c_3 - \mathbf{c}_B^\mathsf{T}\mathbf{B}^{-1}\mathbf{A}_3 = 5/2,\ \bar{c}_5 = c_5 - \mathbf{c}_B^\mathsf{T}\mathbf{B}^{-1}\mathbf{A}_5 = -1/2.$
- $\bar{c}_5 < 0 \implies$ the current solution is not optimal; the nonbasic variable x_5 is to enter the basis.

- $\mathbf{d}_B = -\mathbf{B}^{-1}\mathbf{A}_5 = \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \end{bmatrix}$. The simplex method does not terminate with an unbounded optimum.
- $\theta^* = \min_{i:d_B(i)<0}\{x_{B(i)}/(-d_{B(i)})\} = 4$.
- The basis variable x_4 is to exit the basis.

When $k = 4$,

- $\mathbf{B} = [\mathbf{A}_2, \mathbf{A}_5, \mathbf{A}_1] = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{bmatrix}$, $\mathbf{B}^{-1} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1 & 1 \\ 1/2 & -1/2 & 0 \end{bmatrix}$.
- $\mathbf{x}_B = \begin{bmatrix} x_2 \\ x_5 \\ x_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{x}_N = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
- $\bar{c}_3 = c_3 - \mathbf{c}_B^\intercal \mathbf{B}^{-1}\mathbf{A}_3 = 5/2,\ \bar{c}_4 = c_4 - \mathbf{c}_B^\intercal \mathbf{B}^{-1}\mathbf{A}_4 = 1/2$.
- $\bar{c}_3, \bar{c}_4 \geq 0 \implies$ the current solution is optimal.

The optimal solution of the LP is $(x_1, x_2, x_3, x_4, x_5) = (1, 3, 0, 0, 4)$, and the optimal cost is -11.

4 (25 pt) Modeling exercise: least squares and robust regressions

1. (5 pt) Yes, the objective function of (ADR) is a convex function in $\beta_0, \ldots, \beta_n$. We denote $\beta = (\beta_0, \ldots, \beta_m)$ and the linear functions of β by $f_i(\beta) = y_i - \beta_0 - \sum_{j=1}^n \beta_j x_{ij},\ i = 1, \ldots, N$. Then, the function $g_i = |f_i|$ is convex in β, because for any β^1, β^2 and $\lambda \in [0, 1]$, we have

$$\begin{aligned} g_i(\lambda\beta^1 + (1-\lambda)\beta^2) &= \left|f_i(\lambda\beta^1 + (1-\lambda)\beta^2)\right| \\ &= \left|\lambda f_i(\beta^1) + (1-\lambda) f_i(\beta^2)\right| \\ &\leq \lambda\left|f_i(\beta^1)\right| + (1-\lambda)\left|f_i(\beta^2)\right| \\ &= \lambda g_i(\beta^1) + (1-\lambda) g_i(\beta^2). \end{aligned}$$

   Therefore, the objective function of (ADR), which is the sum of g_i over i, is also convex in β.

2. (5 pt) The problem (ADR) can be formulated as a linear program with auxiliary

variables z_i, $i = 1, \ldots, N$:

$$\min_{z_i, i=1,\ldots,N, \beta_0,\ldots,\beta_n} \sum_{i=1}^{N} z_i,$$

$$\text{s.t. } z_i \geq y_i - \beta_0 - \sum_{j=1}^{n} \beta_j x_{ij},\ i = 1, \ldots, N,$$

$$z_i \geq -(y_i - \beta_0 - \sum_{j=1}^{n} \beta_j x_{ij}),\ i = 1, \ldots, N.$$

3. (10 pt) The solution is that $(\beta_0, \beta_1, \beta_2) = (0.477679, 0.172328, -0.203496)$. The optimal objective value is 25.3586.

4. (5 pt) The plot is shown in Figure 4.

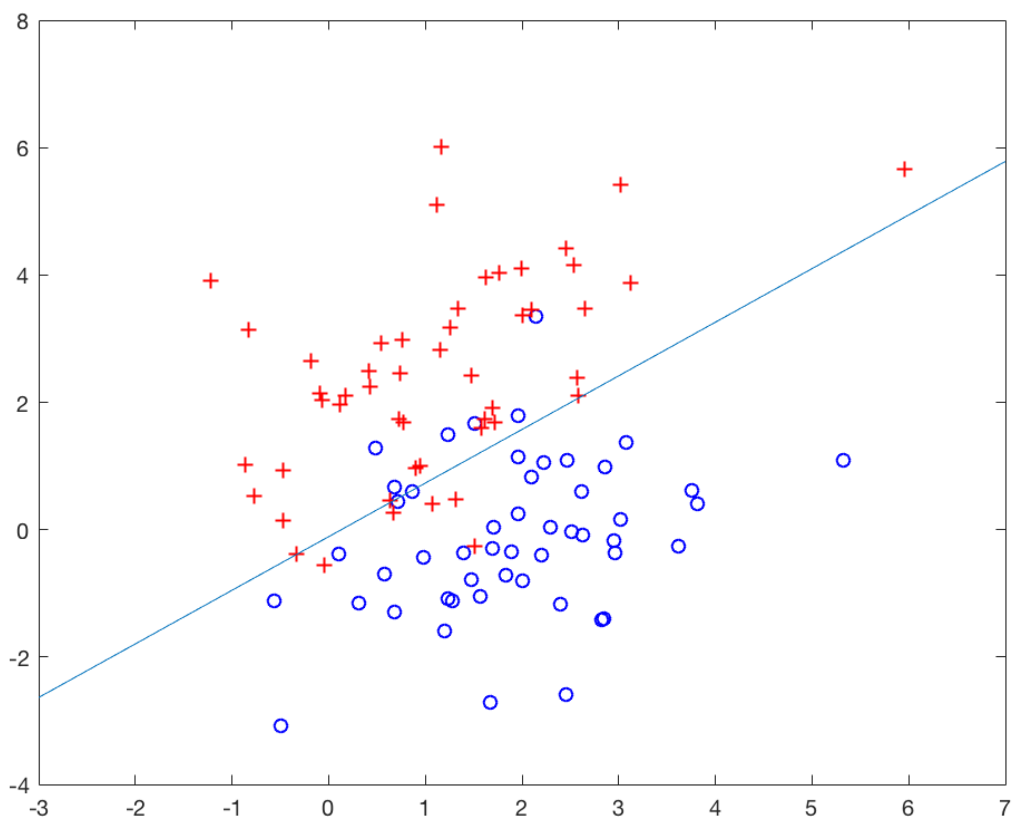
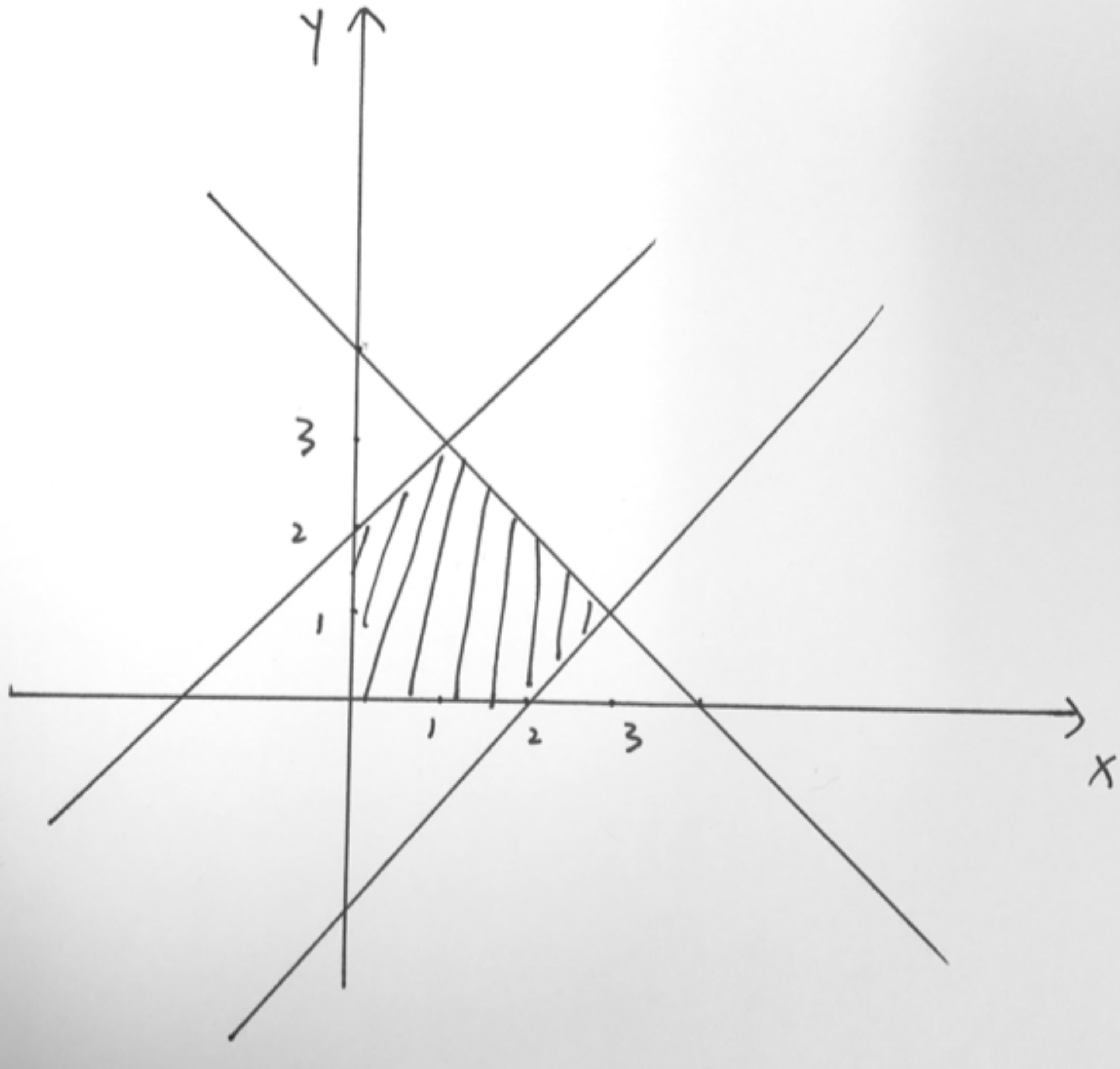


Figure 1: Classification hyperplanes of ADR



Y
X
1
2
3
1
2
3