Optimization

Fall semester 2022/23

Kristóf Bérczi

Eötvös Loránd University
Institute of Mathematics
Department of Operations Research

General information

- **Objective:** To give an overview of basic techniques and results. An ideal outcome is that you can use these ideas in your work after this course.
  - basics of linear programming, LP solving techniques, integer programming, convex sets and functions, convex optimization
- **Course structure: 6** lectures, each lecture consists of a **theoretical** and a **practical** part
- **Course requirement:** 50% exam (end of semester), 50% homework (100%=30pts in total)
- **Contact:** `kristof.berczi@ttk.elte.hu`, Room 3-502
- **Reading:**
  - D. Bertsimas, J.N. Tsitsiklis. Introduction to linear optimization.
  - N. Vishnoi. Algorithms for convex optimization.
  - L.C. Lau. Convexity and optimization.
  - S. Bubeck. Convex Optimization: Algorithms and Complexity.
  - S. Boyd, L. Vandenberghe. Convex Optimization.
  - 'I'm a bandit' blog by Sébastien Bubeck.
  - '3Blue1Brown' channel by Grant Sanderson.

Lecture 1: Linear programming

Systems of linear equations

Example: A firm produces two different goods using two different raw materials. The available amounts of materials are 12 and 5, respectively. The goods require 2 and 3 units of the first material, and both require 1 unit of the second material. Find a production plan that uses all the raw materials.

Idea: Let x_1 and x_2 denote the amounts of the first and second goods produced, respectively. Then the constraints can be written as

$$2 \cdot x_1 + 3 \cdot x_2 = 12$$
$$x_1 + x_2 = 5$$

$$\begin{array}{cc} x_1 & x_2 \end{array}$$
$$\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 12 \\ 5 \end{bmatrix}$$

Solution

Step 1. $x_1 = 5 - x_2$

Step 2. $10 - 2 \cdot x_2 + 3 \cdot x_2 = 12 \Rightarrow x_2 = 2$

Step 3. $x_1 = 5 - 2 = 3$

In general

In general: Gauss elimination

$$\begin{array}{|cccc|}\hline x_1 & x_2 & \dots & x_n \\ \hline \end{array} \qquad \begin{array}{|cccc|}\hline x_1 & x_2 & \dots & x_n \\ \hline \end{array}$$

$$\begin{array}{|cccc|}\hline a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \\ \hline \end{array} = \begin{array}{|c|}\hline b_1 \\ b_2 \\ \vdots \\ b_m \\ \hline \end{array} \Rightarrow \begin{array}{|cccc|}\hline 1 & a'_{12} & \dots & a'_{1n} \\ 0 & 1 & \dots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a'_{mn} \\ \hline \end{array} = \begin{array}{|c|}\hline b'_1 \\ b'_2 \\ \vdots \\ b'_m \\ \hline \end{array}$$

Reduction of the matrix using **elementary row operations**, such as

- swapping two rows,
- multiplying a row by a nonzero number,
- adding a multiple of a row to another row.

Remarks:

- The set of solutions does not change.
- A final solution is ‘easy’ to read out.

Existence of a solution

Assume that your boss gives you such a problem, that is, solve $Ax = b$.

How to prove that a solution exists?

- Just provide a solution x.

How to prove that there is **no** solution?

- Gauss elimination concludes whether there exists a solution or not.
  BUT: this requires the understanding of the algorithm (that you cannot necessarily assume about your boss...)
- Would it be possible to provide some 'shorter' proof?

Fredholm alternative theorem

Fredholm alternative theorem

There exists an x satisfying $Ax = b$ if and only if there exists no y such that $yA = 0,\ yb \neq 0$.

Proof of 'only if' direction.

We show that at most one of x and y may exist. Suppose to the contrary that x and y are such that $Ax = b$ and $yA = 0, yb \neq 0$. Then

$$0 = (yA)x = y(Ax) = yb \neq 0,$$

a contradiction. □

Conclusion: The non-existence of a solution can be proved by providing y.

Geometric interpretation

Naming convention:

Primal problem

$$Ax = b \quad \textbf{(P)}$$

Dual problem

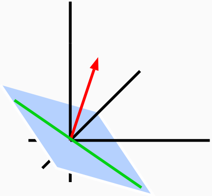
$$\begin{aligned} yA &= 0 \\ yb &\neq 0 \end{aligned} \quad \textbf{(D)}$$

$\Rightarrow$ Fredholm's theorem states that exactly one of **(P)** and **(D)** has a solution.

- The set $H := \{x : ax = b\}$ is a **hyperplane**.
- y is a **normal vector** of the hyperplane $H = \{x : ax = b\}$ if $yx = 0$ for every $x \in H$.

Fredholm as separation theorem

Either b lies in the subspace generated by the columns of A, or it can be separated from it by a homogeneous hyperplane with normal vector y.



The diet problem

What happens if, instead of equalities, a system of linear inequalities is given?

Example: A list of available foods is given together with the nutrient content. Furthermore, the requirement per day of each nutrient is also prescribed. For example, the data corresponding to two types of fruits (F1 and F2) and three types of nutrients (fats, proteins, vitamins) is as follows:

	Fats	Proteins	Vitamins	Available
F1	1	4	5	3
F2	0	2	9	5
Req.	1	5	14	

The problem is to find how much of each fruit to consume per day so as to get the required amount per day of each nutrient, if one can consume at most 2 kg of fruits per day.

Modeling the problem

	Fats	Proteins	Vitamins	Available
F1	1	4	5	3
F2	0	2	9	5
Req.	1	5	14	

Let x_1 and x_2 denote the amounts of fruits F1 and F2 to be consumed per day.

$$\begin{aligned} x_1 \qquad\quad &\geqslant 1 \\ 4x_1 + 2x_2 &\geqslant 5 \\ 5x_1 + 9x_2 &\geqslant 14 \\ x_1 + \ x_2 &\leqslant 2 \end{aligned}$$

Questions:

- How to decide feasibility?
- How to find a solution (if exists) algorithmically?
- How to verify that there is no solution?

Different forms

Observations:

- An equality $ax = b$ can be represented as a pair of inequalities $ax \leqslant b$ and $-ax \leqslant -b$.
- An inequality $ax \leqslant b$ can be represented as the combination of an equality $ax + s = b$ and a non-negativity constraint $s \geqslant 0$, where s is called a **slack** variable.
- A non-positivity constraint $x \leqslant 0$ can be expressed as a non-negativity constraint $-x \geqslant 0$.
- A variable x unrestricted in sign can be replaced everywhere by $x^+ - x^-$, where $x^+, x^- \geqslant 0$.

General form

$$Px_0 + Ax_1 = b_0$$
$$Qx_0 + Bx_1 \leqslant b_1$$
$$x_1 \geqslant 0$$

Standard form

$$Ax = b$$
$$x \geqslant 0$$

Canonical form

$$Qx \leqslant b$$
$$x \geqslant 0$$

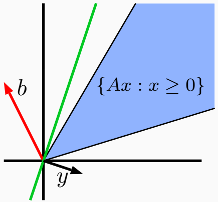
Farkas' lemma

Farkas' lemma, standard form

There exists an x satisfying **(P)** $Ax = b, x \geqslant 0$ if and only if there exists no y such that **(D)** $yA \geqslant 0, yb < 0$.

Farkas' lemma as separation theorem

Either b lies in the cone generated by the columns of A, or it can be separated from it by a homogeneous hyperplane with normal vector y.



Proof of the 'only if' direction.

We show that at most one of x and y may exist. Suppose to the contrary that x and y are such that $Ax = b, x \geqslant 0$ and $yA \geqslant 0, yb < 0$. Then

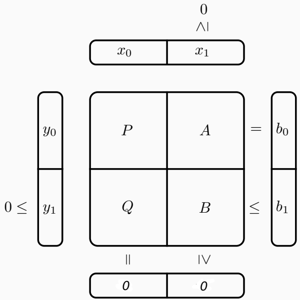
$$0 \leqslant (yA)x = y(Ax) = yb < 0,$$

a contradiction. □

Farkas' lemma in general

Farkas' lemma, general form

There exists an $x = (x_0, x_1)$ satisfying
(P) $Px_0 + Ax_1 = b_0, Qx_0 + Bx_1 \leqslant b_1, x_1 \geqslant 0$
if and only if there exists no $y = (y_0, y_1)$ such that
(D) $y_0 P + y_1 Q = 0, y_0 A + y_1 B \geqslant 0, y_1 \geqslant 0, y_0 b_0 + y_1 b_1 < 0.$



Conclusion: The feasibility/infeasibility of a system of linear inequalities can be proved by providing a solution to the primal/dual problem, respectively.

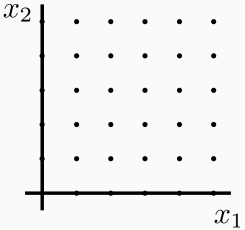
Remaining question: How to find such a solution?

$\Rightarrow$ We will answer this in a far more general setting!

Geometric background I

Example

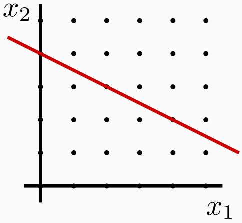
$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



Geometric background I

Example

$$x_1 + 2 \cdot x_2 \leqslant 8$$
$$2 \cdot x_1 + x_2 \leqslant 6$$
$$x_1, x_2 \geqslant 0$$



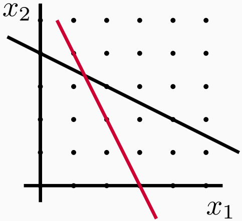
Geometric background I

Example

$$x_1 + 2 \cdot x_2 \leqslant 8$$

$$2 \cdot x_1 + x_2 \leqslant 6$$

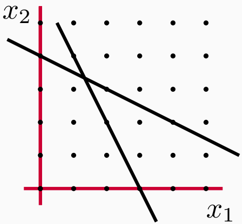
$$x_1, x_2 \geqslant 0$$



Geometric background I

Example

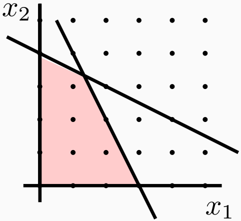
$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



Geometric background I

Example

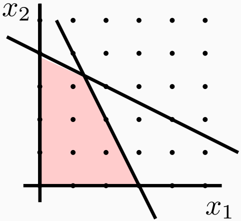
$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



Geometric background I

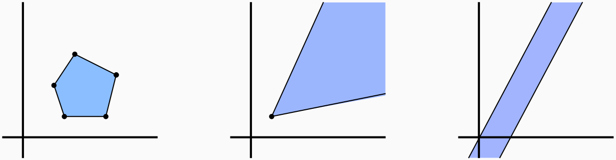
Example

$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



- An inequality $ax \leqslant b$ defines a **half space**.
- The solution set is the **intersection** of a finit number of half spaces, called a **polyhedron**.

Geometric background II



- Given a polyhedron P, a point $x \in P$ is a **vertex** of P if there exists no y such that $x + y, x - y \in P$.
- A **polytope** is the convex hull of a finite number of points.

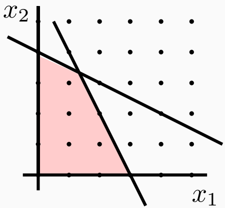
Thm.

Every polytope is a polyhedron, and every bounded polyhedron is the convex hull of its vertices.

Geometric background III

Example

$$
\begin{aligned}
x_1 + 2 \cdot x_2 &\leqslant 8 \\
2 \cdot x_1 + \quad x_2 &\leqslant 6 \\
x_1, x_2 &\geqslant 0
\end{aligned}
$$



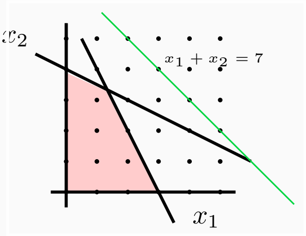
Goal: Maximize/minimize a linear objective function over the set of solutions.

$\Rightarrow$ Example: $\max\{x_1 + x_2\}$.

Geometric background III

Example

$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



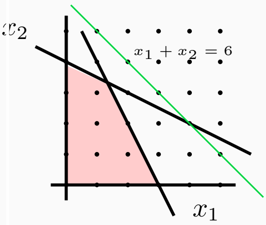
Goal: Maximize/minimize a linear objective function over the set of solutions.

$\Rightarrow$ Example: $\max\{x_1 + x_2\}$.

Geometric background III

Example

$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



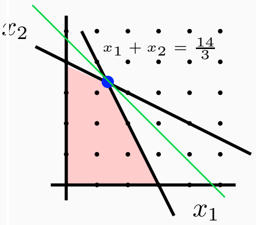
Goal: Maximize/minimize a linear objective function over the set of solutions.

⇒ Example: $\max\{x_1 + x_2\}$.

Geometric background III

Example

$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



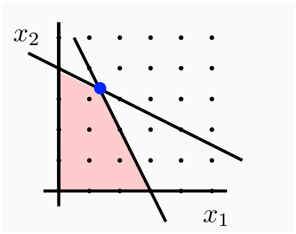
Goal: Maximize/minimize a linear objective function over the set of solutions.

⇒ Example: $\max\{x_1 + x_2\}$.

Geometric background III

Example

$$\begin{aligned} x_1 + 2 \cdot x_2 &\leqslant 8 \\ 2 \cdot x_1 + \quad x_2 &\leqslant 6 \\ x_1, x_2 &\geqslant 0 \end{aligned}$$



Goal: Maximize/minimize a linear objective function over the set of solutions.

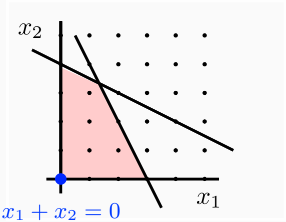
$\Rightarrow$ Example: $\max\{x_1 + x_2\}$.

Idea: Start from a vertex, and move to a neighboring vertex with improved objective value.

Geometric background III

Example

$$x_1 + 2 \cdot x_2 \leqslant 8$$
$$2 \cdot x_1 + x_2 \leqslant 6$$
$$x_1, x_2 \geqslant 0$$



Goal: Maximize/minimize a linear objective function over the set of solutions.

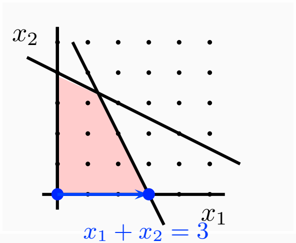
⇒ Example: $\max\{x_1 + x_2\}$.

Idea: Start from a vertex, and move to a neighboring vertex with improved objective value.

Geometric background III

Example

$$x_1 + 2 \cdot x_2 \leqslant 8$$
$$2 \cdot x_1 + x_2 \leqslant 6$$
$$x_1, x_2 \geqslant 0$$



Goal: Maximize/minimize a linear objective function over the set of solutions.

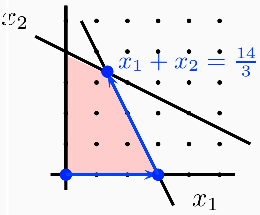
$\Rightarrow$ Example: $\max\{x_1 + x_2\}$.

Idea: Start from a vertex, and move to a neighboring vertex with improved objective value.

Geometric background III

Example

$$
\begin{aligned}
x_1 + 2 \cdot x_2 &\leqslant 8 \\
2 \cdot x_1 + \quad x_2 &\leqslant 6 \\
x_1, x_2 &\geqslant 0
\end{aligned}
$$



Goal: Maximize/minimize a linear objective function over the set of solutions.

$\Rightarrow$ Example: $\max\{x_1 + x_2\}$.

Idea: Start from a vertex, and move to a neighboring vertex with improved objective value.

History

1827, Fourier: Fourier-Motzkin elimination
1939, Kantorovich: reducing costs of army, general LP
1940's, Koopmans: economic problems as LPs
1941, Hitchcock: transportation problems as LPs
1946-47, Dantzig: general LP for planning problems in US Air Force (simplex method)
1979, Khachiyan: ellipsoid method, LP is solvable in linear time (more theoretical than practical)
1984, Karmakar: interior-point method (can be used in practice)

Linear programs

We would like to solve problems of the form

General form

$$\begin{aligned} \max \ & c_0 x_0 + c_1 x_1 \\ \text{s.t. } & P x_0 + A x_1 = b_0 \\ & Q x_0 + B x_1 \leqslant b_1 \\ & x_1 \geqslant 0 \end{aligned}$$

Standard form

$$\begin{aligned} \max \ & cx \\ \text{s.t. } & Ax = b \\ & x \geqslant 0 \end{aligned}$$

Canonical form

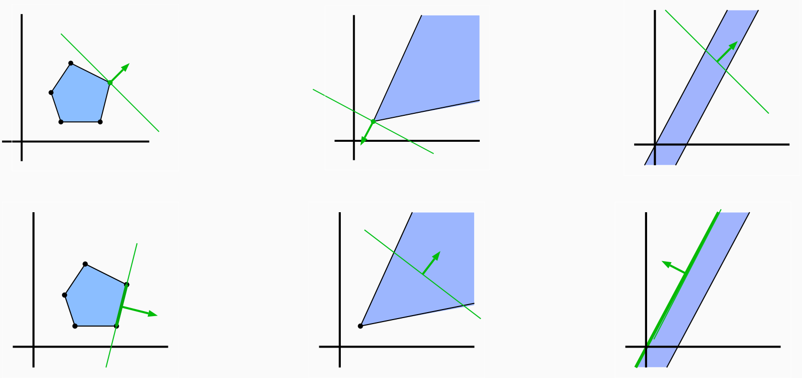
$$\begin{aligned} \max \ & cx \\ \text{s.t. } & Qx \leqslant b \\ & x \geqslant 0 \end{aligned}$$

Remarks:

- A minimization problem min cx can be reformulated as a maximization problem max $(-c)x$ and vice versa.
- The optimal solution can be obtained by 'moving' a hyperplane with normal vector c towards the polyhedron, and finding the first point where they meet [**Be careful:** min or max?]

  $\Rightarrow$ Intuition: the optimum is always attained at a vertex.

Geometric background IV



Possible cases:

- single optimal solution,
- infinite number of optimal solutions, or
- no optimal solution (unbounded objective value).

Geometric background V

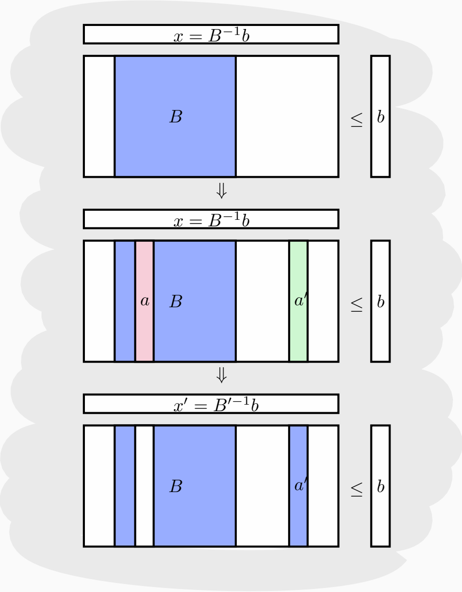
Thm.

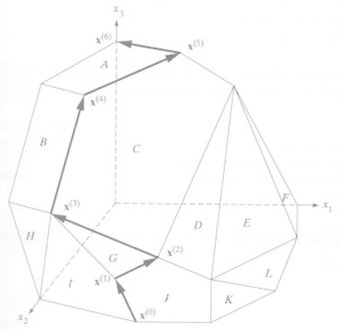
Let $P = \{x : Qx \leqslant b\}$ where the columns of Q are linearly independent. Then $x \in P$ is a vertex if and only if it can be obtained by taking a non-singular $r(Q) \times r(Q)$ submatrix Q' of Q and the corresponding part b' of b, and solving the system $Q'x = b'$.

Remarks:

- The number of such submatrices, and so the number of vertices is finite.
  $\Rightarrow$ If each vertex is visited at most once, then the procedure terminates.
- When the columns are non-independent, then there is an infinite number of basic feasible solutions. However, there are only a finite number of so-called **strong basic feasible solutions**, and, if it exists, the optimum is attained in one of them.

Simplex method





Problems

- Running time?
- Optimal solution?

Running time

Problem 1: The simplex algorithm might fail to terminate.

Reason: The algorithm can fall into cycles between bases associated with the same basic feasible solution and objective value.

Solution: Careful pivoting rule, e.g. **Bland's rule** prevents cycling.

Problem 2: Efficient in practice, but for almost every variant, there is a family of linear programs for which it performs **badly**.

Reason: The number of vertices of a polyhedron can be exponentially large.

Solution: Sub-exponential pivot rules are known.

Major open problem: Is there a variant with polynomial running time?

- **Hirsch's conjecture:** Let P be a d-dimensional convex polytope with n facets. Then the diameter of P is at most $n - d$.
- Counterexample by Francisco Santos, 2011 (86 facets, 43-dimensional).

Duality theorem

Problem 3: Is the solution optimal?

Duality theorem

Consider the problems

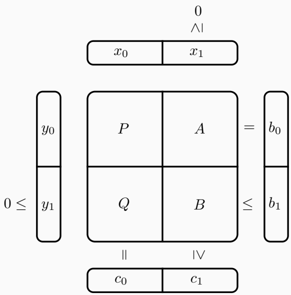
(P) $\max(c_0 x_0 + c_1 x_1)$ s.t. $P x_0 + A x_1 = b_0, Q x_0 + B x_1 \leqslant b_1, x_1 \geqslant 0$

and

(D) $\min(y_0 b_0 + y_1 b_1)$ s.t. $y_0 P + y_1 Q = c_0, y_0 A + y_1 B \geqslant c_1, y_1 \geqslant 0$.

Then exactly one of the followings hold:

1. both (P) and (D) are empty,
2. (D) is empty and (P) is unbounded,
3. (P) is empty and (D) is unbounded,
4. both (P) and (D) have a solution, and $\max = \min$.



Reading assignment

- D. Bertsimas, J.N. Tsitsiklis. Introduction to linear optimization.
  - Chapter 1, Sections 1.1, 1.2, 1.4, and 1.5
  - Chapter 2, Sections 2.1 and 2.2
  - Chapter 4, Sections 4.1-4.3, 4.6
- A. Frank. Operációkutatás (in Hungarian).
  - Chapter 2
  - Chapter 3
  - Chapter 4

Exercises

Submission deadline: The starting time of the next lecture.

1. Bob would like to write down the system $3x + 2y + 4z = 8$, $-3y \leqslant 3$, $x - 3z \geqslant 10$, $\min x - y$, but his keyboard is missing the symbols $=$ and $\geqslant$, and the letter i is not working. Reformulate the problem only using $\leqslant$ and maximization. (1pt)
2. Prove that the system $Ax \leqslant 0, x \gg 0$ admits a solution if and only if $Ax \leqslant 0, x \geqslant 1$ has one. (1pt)
3. Consider the problem $x_2 \leqslant 4, x_1 + x_2 \leqslant 6, 2x_1 + x_2 \leqslant 10, x_1, x_2 \geqslant 0$. Represent these constraints on the plane. Find a point that maximizes $x_1 + 2x_2$. (2pts)
4. Verify the 'only if' direction in the general form of Farkas' lemma. (1pt)
5. Assume that both (P) and (D) has a solution in the duality theorem. Prove that weak duality holds, that is, max $\leqslant$ min. (1pt)
6. Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c_1, \ldots, c_k \in \mathbb{R}^n$. Formulate the following problem as an LP: $Ax = b$, $x \geqslant 0$, $\min f(x)$, where $f(x) := \max\{c_1 x, \ldots, c_k x\}$. (1pt)
7. Reduce the following systems of inequalities to each other (in the sense that if we can solve one of them, then we can solve any of them):

   $$\begin{matrix} Ax = b \\ x \geqslant 0 \end{matrix} \qquad \begin{matrix} Bx \leqslant b \\ x \geqslant 0 \end{matrix} \qquad Qx \leqslant b \qquad \begin{matrix} Px_0 = b_0 \\ Qx \leqslant b_1 \end{matrix}$$

   Write up Farkas' lemma for all of them. (3pts)