Parabola Volume 59, Issue 3 (2023)

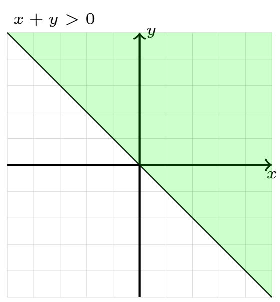
The role of straight line coefficients in linear programming

Hoover H. F. Yin[1]

1 Introduction

The graphical method [1, 2] is an excellent tool for visualising the basic concept of linear programming in secondary school. After plotting the straight lines associated with the constraints, we need to identify which polyhedron is the feasible region, i.e., the region that all constraints are satisfied. Usually, students are taught to substitute an arbitrary point to each constraint and see whether the strict inequality holds. This way, the incorrect half-plane of each constraint can be eliminated.

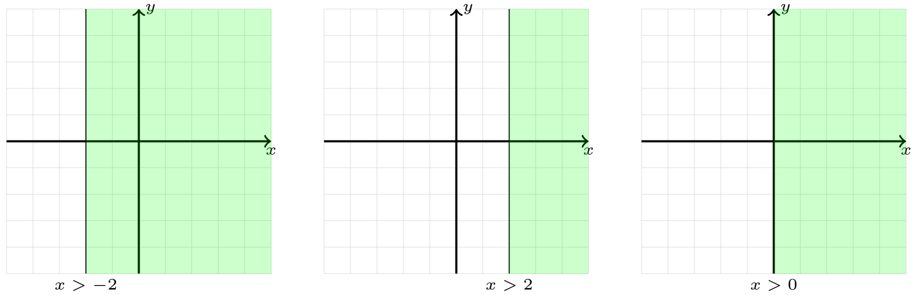
As a quick illustration, suppose we have the constraint $x+y>0$. The line $x+y=0$ is plotted in the figure below.



Which one is the correct half-plane? The green or the white one? A common technique is to substitute $(0,0)$ to the constraint so that the variables x and y can be eliminated. However, $(0,0)$ lies on the line in this example thus we cannot draw any conclusion if we do so. We need to consider another point, say, $(1,1)$, which is on the green region. Then, we obtain a valid strict inequality $2>0$. This suggests that the green one is the correct half-plane.

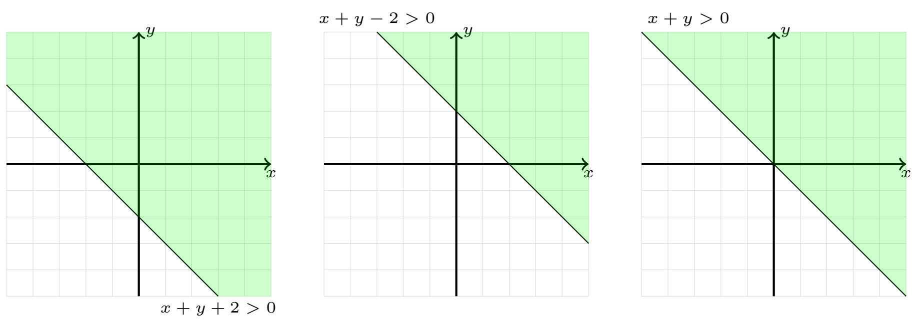
However, when the line is a vertical (or horizontal) one, we can identify the correct half-plane instantaneously without substitution. For example, it is trivial that the green region in each of the following plots is the correct half-plane.

[1]Hoover H. F. Yin is a lecturer at the Department of Information Engineering, The Chinese University of Hong Kong.



The coefficient of x is positive, thus we have the intuition (from real number line) that "greater than" means the right side and "less than" means the left side.

How about sloped lines? Let us consider the following examples.



Again, the green regions are the correct half-planes. Although the origin is not located on the same side of the above lines, the correct half-planes are all on the right side. One can try more parallel lines to see that no matter what the y-intercept is, the correct half-plane is always on the right side.

This phenomenon looks like that when the coefficient of x is positive: "greater than" means the direction pointing to $+\infty$ (the right side) and "less than" means the direction pointing to $-\infty$ (the left side). With this observation, it is natural to ask: Is there any simple way to instantaneously identify the correct half-plane **without** substitution? The answer is affirmative: The coefficients of the straight line tell us everything!

2 The role of straight line coefficients

We can view the problem from two different perspectives: The algebraic point of view, and the geometric point of view. Before I explain further, let me first give the answer to our question.

Consider a constraint

$$Ax + By + C > 0.$$

Note that:

(a) We have the same correct half-plane (but including the line) if we instead consider $Ax + By + C \geq 0$.

(b) When the inequality sign of the constraint is reversed, we need to reverse the decision in the following rules (equivalently, multiply -1 to both sides to flip the inequality sign).

The role of A, the coefficient of x:

- If $A > 0$, then the right side (the $+\infty$ side of the x-axis) is the correct half-plane.
- If $A < 0$, then the left side (the $-\infty$ side of the x-axis) is the correct half-plane.

The role of B, the coefficient of y:

- If $B > 0$, then the top side (the $+\infty$ side of the y-axis) is the correct half-plane.
- If $B < 0$, then the bottom side (the $-\infty$ side of the y-axis) is the correct half-plane.

The role of C, the constant term:

- If $C > 0$, then the side which contains the origin is the correct half-plane.
- If $C < 0$, then the side which does not contain the origin is the correct half-plane.

Algebraic explanation

The role of C is easy to explain. By substituting $(0, 0)$ to the constraint, the inequality becomes $C > 0$. If it is true, then the side containing the origin is the correct half-plane, and vice versa.

To explain the roles of A and B, we first go back to basics: Why can we identify the correct half-plane for any vertical line $x > c$ instantaneously? A fundamental reason is that we know the existence of c' such that $c' > c$, say, $c' = c+1$. By substituting the point $(c', 0)$, we obtain a valid inequality so the right half-plane is the correct region. That is, we are choosing a point to substitute that can guarantee a valid strict inequality.

To proceed, we reorder the terms of the constraint and obtain

$$Ax > -By - C. \tag{1}$$

If $A > 0$, then we substitute a point (x', y) where $x' > \frac{-By-C}{A}$. We have the freedom to select y, so for simplicity, we choose $y = 0$. Then, our desired point is $(x', 0)$ where $x' > \frac{-C}{A}$. We know the existence of x' such that $x' > \frac{-C}{A}$, so we know that the correct half-plane is on the right side. We can alternatively say that we substitute a point $(x, 0)$ where x tends to infinity. By doing so, we have $\lim_{x\to+\infty} Ax + By + C = +\infty$, which gives the same conclusion.

If $A < 0$, then the result is reversed because the inequality sign is flipped when we divide both sides of (1) by A. By a similar argument, we know that the correct half-plane is on the left side.

We can also explain the role of B in a similar manner. On the other hand, it is not hard to extend this argument to a higher-dimensional space for a constraint involving more than two variables.

Geometric explanation

As we can instantaneously identify the correct half-plane when the line is vertical, how about if we rotate the axes to make the sloped line vertical? For a line $Ax+By+C=0$, we can write it as

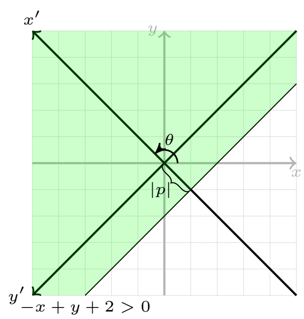
$$\frac{A}{\sqrt{A^2+B^2}}x+\frac{B}{\sqrt{A^2+B^2}}y-\frac{-C}{\sqrt{A^2+B^2}}=0,$$

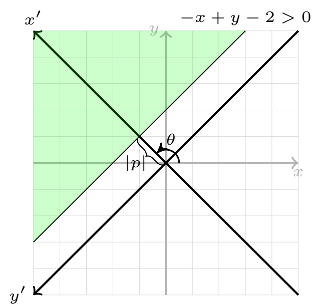
which matches with the normal form of straight line

$$x\cos\theta+y\sin\theta-p=0,$$

where p (possibly negative!) is the length of the normal vector drawn from the origin to the line which subtends an angle θ with the positive direction of the x-axis. As $\sqrt{A^2+B^2}>0$, we know that the constraint $Ax+By+C>0$ is equivalent to $x\cos\theta+y\sin\theta-p>0$.

The figures below illustrate two examples of rotated axes.
The left figure considers the constraint $-x+y+2>0$, where $\theta=\frac{3\pi}{4}$ and $p=-\sqrt{2}$.
The right figure considers the constraint $-x+y-2>0$, where $\theta=\frac{3\pi}{4}$ and $p=\sqrt{2}$.





In the above examples, the $x'y'$-plane is obtained by rotating the xy-plane counterclockwise through an angle θ about the origin. In other words, every point on the xy-plane is rotated clockwise through an angle θ about the origin to become a point on the $x'y'$-plane. By applying the rotation matrix, we have

$$\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\cos\theta+y\sin\theta\\-x\sin\theta+y\cos\theta\end{pmatrix}.$$

Then, the constraint $x\cos\theta+y\sin\theta-p>0$ becomes $x'>p$, which means that the correct half-plane is the one where the x'-axis points to the positive infinity.

One may easily confuse the angle θ with $\theta+\pi$. A way to find the correct θ is to check the sign of p. A positive p means that the line is on the positive side of the x'-axis, so the correct half-plane does not contain the origin, and vice versa. The sign of p is opposite to the sign of C, so the interpretation of p explains the role of C.

The role of A can be interpreted this way. As long as we know which side the x'-axis points to the positive infinity, we are done. Note that the signs of A and $\cos\theta$ are the same. If $A > 0$, then we know that the correct θ is in the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$, i.e., $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$, so the x'-axis points to the right side with a certain acute angle. If $A < 0$, then θ is in the interval $(\frac{\pi}{2}, \frac{3\pi}{2})$, i.e., $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$, which corresponds to the left side.

The role of B can be explained similarly. The signs of B and $\sin\theta$ are the same. If $B > 0$, then θ is in the interval $(0, \pi)$, i.e., $0 < \theta < \pi$, which is the top side. If $B < 0$, then θ is in the interval $(\pi, 2\pi)$, i.e., $\pi < \theta < 2\pi$, which is the bottom side.

3 Conclusion

The graphical method is a useful tool to introduce linear programming. Substituting points in the constraints is a standard method taught to identify feasible regions. This article describes a pedagogical observation about the relation between the straight line coefficients in a constraint and the correct half-plane. With this observation, we can identify the correct half-plane without numerically substituting points! I hope that this article can inspire students to ask questions, think outside the box, and help them appreciate the linkage between different topics in mathematics.

References

[1] J.E. Reeb and S.A. Leavengood, Using the graphical method to solve linear programs, EM 8719, Oregon State University, USA, 1998.

[2] P. Semmes, A geometric approach to linear programming in the two-year college, *The Two-Year Coll. Math. J.* **5** (1974), 37–40.