Linear Programming and Its Mathematical Roots

Roger GODARD

Département de mathématiques et d'informatique, Collège militaire royal du Canada, Kingston (On.) Canada

godard-r@rmc.ca

Abstract

Several important aspects of Linear Programming are reviewed and commented: (1) the geometric aspect and convexity, (2) the duality concept, (3) the sensitivity analysis on variables and coefficients, (4) the links with Linear Algebra and systems of inequalities, and finally (5) the algorithms.

1 Introduction to linear programming and optimization problems

In his book "Linear Programming and Extensions", Dantzig (1963) presented a table to trace back its History. Our intention is to perform a traveling on the roots of Linear Programming and on its multidisciplinary aspects by using Dantzig's references but with further emphasis on the development of the mathematical tools.

Several important aspects of Linear Programming have been neglected in former studies on the origins of Linear Programming: (1) the geometric aspect and convexity, (2) the duality concept, (3) the sensitivity analysis on variables and coefficients, (4) the links with Linear Algebra, (5) and the algorithms.

If we scan speedily the history of the optimization methods, we remember Lagrange's multipliers method for the optimization of constrained problems. Lagrange published his essay in 1762, and also, in his "Théorie des fonctions analytiques" in 1797. After Cauchy, who, in 1827, made the first application of the steepest descent method to solve unconstrained minimization problems, we observe very little progress made afterwards until the middle of the twentieth century. Dantzig's table (1963) had given some key dates for the development of linear programming, and some associated optimization methods. The development of linear programming is mainly associated to such names as Kantorovich (1939) and Dantzig, in 1947. Then, in 1951, Kuhn and Tucker provided the necessary and sufficient conditions of optimality in non-linear programming.

2 Linear programming: objective function and linear inequalities

A linear programming problem is to minimize a linear objective function $f(x) = c^t x$ subject to a set of linear constraints $Ax = b$; $x \geq 0$. These constraints may be equality or inequality constraints. In the latter case, an inequality constraint can be converted to an equality constraint by introducing a positive (negative) variable which is called a slack variable. These constraints are hyperplanes, and the set of solutions is a convex polyhedron. Then, at least, one of the vertices of the convex polytope should correspond to the optimum solution. Therefore the simplex or Dantzig algorithm was to compare the solutions at vertices in an orderly way in order to find an optimal path towards the true solution. Figure 1 is taken from Kantorovich's 1939 article. It illustrates the feasibility convex region for a transportation problem:

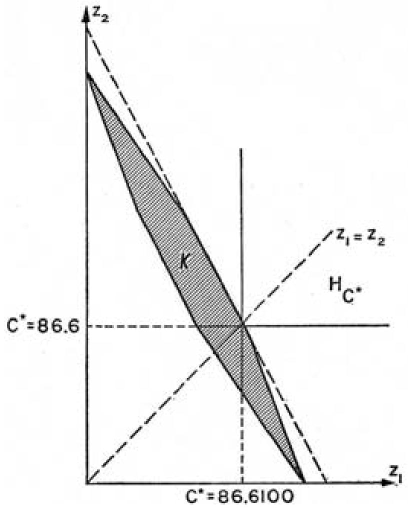

Figure 1 – Feasibility region for the best plan of freight shipments (Kantorovich, 1939)

3 Slack variables and solving systems of inequalities

In 1798, while he was working on problems of statics, J. Fourier had to solve systems of linear inequalities. Again Fourier published on that particular topic in 1823, 1824, and 1826. He then suggested that a theory of systems of such inequalities should be developed. He even proposed that his method could be used in Geometry, Algebraic Analysis, Mechanics (Statics), and Theory of Probability. Most probably, when he refereed on the theory of probability, he had in mind the theory of errors in sciences of observations: "Donner au plus grand écart, sa moindre valeur" i.e. a minimization process in the ℓ_∞ norm. As for the solution of his system of inequalities, Fourier described an elimination method by reducing the number of variables, and a geometrical approach. From six inequalities and in the case of two variables, Fourier built a convex polygon 123456 of the set of feasible solutions:

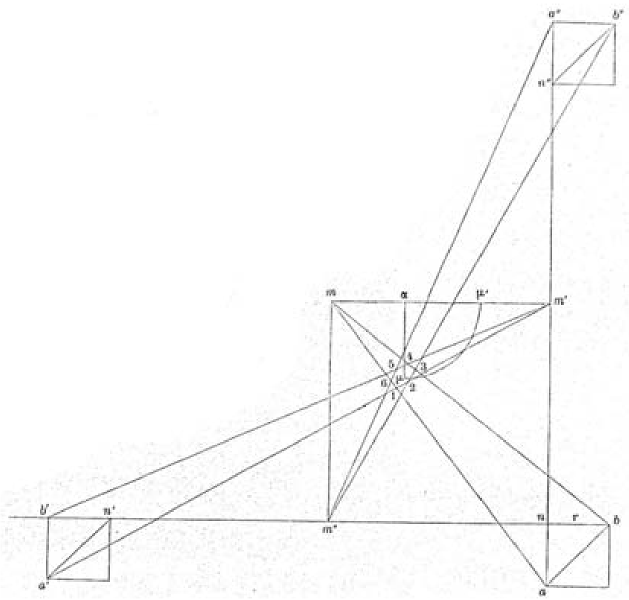

Figure 2 – Fourier's polygon of solutions

« Il faut remarquer que le système de tous ces plans (from the inequalities) forme un vase qui leur sert de limite ou d'enveloppe. La figure de ce vase extrême est celle d'un polyèdre dont la convexité est tournée vers le plan horizontal. » If the inequality decreased, the polygon shrank towards a single point, the center of gravity. The next figure represents the famous polyhedron of feasible solutions.

Linked to the problem of minimization in ℓ_∞ norm is C. de la Vallée Poussin's contribution (1911). Again, he (VP) searched for a solution of an over-determined system of equations, with applications to sciences of observations and the theory of errors. His paper can be considered as a complement to his 1908 article on interpolation formulas (de la Vallée Poussin theorem), but his approach to the minimization of the absolute value of the largest error could be dangerous and sensitive to outliers. VP searched for a pure algebraic approach. He introduced slack variables for residuals. He then selected the equations with the worse residuals (by trial and errors) and minimized these residuals. By selecting his equations, he was able to solve square systems of linear equations with the technique of determinants. In parallel, the minimization in the ℓ_1 norm has always represented a more difficult problem. The first attempts came from Boscovich in 1750, and Laplace in 1786. The first representation of the ℓ_1 problem as a linear programming problem arose in 1955.

The main theoretical contributions to the theory of systems of linear inequalities came from Germany and Eastern Europe. Paul Albert Gordan was born in Breslau, Germany (now Wroclaw, Poland) and died in 1912 in Erlangen, Germany (1837–1912). Published in 1873, his theorem may be formulated as follows: in addition to the system of inequalities $Ax > 0$ one considers the system of equations: $A^t y = 0$; $y \geq 0$, $y \neq 0$. One of the two systems has a solution. We should also mention that Gordan's only doctoral student was Emmy Noether.

We shall now comment on another Dantzig's reference on Minkowski. In 1896, C. Hermite, the French Mathematician, after receiving Minkowski's book Geometrie der Zahlen, wrote: "Je crois voir la terre promise"! The seven pages of paragraph 19, chapter 1 exposed his work on systems of linear inequalities, Minkowski proved that there are finitely many "extreme" solutions, the vertices, such that every solution is a linear combination of these. He also introduced the concept of "slack" variables. These slack variables became a paradigm in LP. They introduced the precious notion of scarcity in a matrix system; moreover, they established a method of communication between the different equations, and the utilisation of vector spaces. They enabled to build a method just as important than Legendre's contribution for the least squares method.

Julius Farkas was a Hungarian, born in 1847, who died in 1930. He was a physicist who also did work on Mathematics, remembered for his 1902 theorem on inequalities. Inspired by Fourier, his 27 pages article is more detailed than Minkowski's paragraph on inequalities. Finally, in 1936, Motzkin's thesis provided the most comprehensive treatment of systems of linear inequalities. Also, at the same time, Mathematicians such as L. L. Dines were interested by convex hulls and linear inequalities, and the search of necessary and sufficient conditions for the existence of a solution of a system of inequalities, and the duality.

4 Convexity

To the algebraic system of inequalities will correspond a geometric interpretation, in terms of convex bodies. We already mentioned that LP constraints are hyperplanes, and the set of solutions is a convex polyhedron. Then, at least, one of the vertices of the convex polytope should correspond to the optimum solution. Therefore, it seems appropriate to review the theory of convexity.

Convexity is an inter-disciplinary, heterogeneous field, which has two branches: geometry and analysis (sets and functions). Notions of convexity probably came first from observations of nature: crystals, stones, trees, with the development of geometric figures such as circles, squares, rectangles, cylinders, etc. One could quote the Pythagoricians with the regular polytopes, Euclide and Archimedes. In his treatise "On Spheres and Cylinders", Archimedes defines a "convex arc as a plane curve which lies on one side of the line joining its endpoints and all chords of which lies on the same side of it."

From the XVII$^{\text{th}}$ and the XVIII$^{\text{th}}$ centuries, we can distinguish two main paths on convexity. One is linked with Descartes, Leibnitz and Euler and the theory of polyhedra. And the other one is linked with the theory of functions and the variational calculus. We shall first consider the problem of convex bodies.

In 1750 and 1751, Leonard Euler made a definitive contribution on the theory of polyhedra, because of the generalizations that occurred and the evolution of ideas in combinatorial topology. His theorem, even if it was stated incorrectly:

"In every solid enclosed by plane faces, the number of faces along with the number of solid angles exceeds the number of edges by two", has the form: $F - E + V = 2$, where F, E, and V denote the number of faces, edges, and vertices of a polyhedron. It was an early example of the problem of a convex body, although, implicitly stated. Euler's formula was known to Descartes around 1630. But this formula provoked many investigations with Legendre, Cauchy, l'Huilier, Gergonne, von Staudt, Steiner, Schläfli, Poinsot, Hessel, Möbius, Listing, Jordan, Poincaré and H. Hopf, P. Alexandrov, etc. The word "simplex" was probably introduced in the mathematical vocabulary by Poincaré. Steinitz (1916) defined a simplex as a bounded convex portion of the Euclidian space determined by $(n + 1)$ linearly independent points. Even if all these prestigious mathematicians did not contribute directly to the field of optimization or LP problems, they had an indirect influence on the study of convex bodies, and the principle of duality, so important in LP. Even more, Albert W. Tucker, the Princeton mathematician, (1905–1995) began his career as a topologist.

Linked to the development of the set theory, convex sets were properly defined by Minkowski and Brunn. David Hilbert who was very close to Minkowski, wrote these following sentences:

> Ein konvexer (nirgends konkaver) Körper ist nach Minkowski als ein solcher Kôrper definiert, der die Eigenschaft hat, dass, wenn man zwei seiner punkte ins Auge fasst, auch die ganze geraldlinige Strecke zwischen denselben zu dem Körper gehört.

Minkowski and after the Gottingen group, made some definitive contributions to the field of convex bodies, their direct sums, intersections of convex sets, convex hulls, etc., where Caratheodory theorem, in 1911 and Eduard Helly theorem, in 1923 which would later have some important applications in LP. Minkowski's book "Geometrie der Zahlen" was published in 1896, and reedited in 1910. And his 1911 "Theorie der konvexen Körper" was an important contribution to the theory of convex cones. We see the emergence of a link between systems of linear inequalities and convex sets or projective geometry. Several proofs were based on 1913–1915 Steinitz's ideas. Convexity appeared a mature mathematical subject in other books such as the one from Bonnesen and Fenchel's 1934 "Theorie der konvexen Körper", W. Fenchel got his first academic position in Göttingen. He later had to escape from the Nazis and went to Copenhagen. In 1951, he lectured on convex sets, and functions at Princeton University, at the time where the Princeton group was leading in linear programming.

Also, at the end of the XIX$^{\text{th}}$ century, convex functions resurfaced in the mathematical aura. An example of this, were the properties of the Euler Gamma function. Some desired fundamental geometrical properties of functions were found in the notion of convexity. In the search for sufficient conditions to a maximum or a minimum, we are conducted to a class of concave or convex functions and a class of convex sets. Independently O. Hölder, in 1889 in

Göttingen and Jensen, in 1906, in Copenhagen gave formal definitions for convex functions. For them, a real, finite and continuous function $f(x)$ of a real variable x, is a convex function in a given interval if the following inequality is true:

$$\frac{f(x)+f(y)}{2} \geq f\left(\frac{x+y}{2}\right)$$

Hölder (1859–1937) used directly a more general definition for convexity:

$$\frac{a_1\varphi(x_1)+a_2\varphi(x_2)+\ldots+a_n\varphi(x_n)}{a_1+a_2+\ldots+a_n} > \varphi\left(\frac{a_1x_1+a_2x_2+\ldots+a_nx_n}{a_1+a_2+\ldots+a_n}\right)$$

The above inequality expressed the relation that a function value at the weighted average of the x_j is not greater than the weighted average of the function values at the x_j. Indeed, mathematical programming was directly concerned with the existence and uniqueness of solutions. And clearly, convex problems on convex sets did guarantee global extrema. A local minimum is also the global minimum. If we take the example of Beckenbach's article, in 1948, convexity was linked to the second derivative of a function, and $f(x)$ is concave if and only if $-f(x)$ is convex. At least two important books, one by W. Fenchel in 1953, and the other one, by H. G. Eggleston in 1958 introduced the differential conditions for convexity. They also provided historical notes and an extensive list of references on convexity.

5 Duality

One of the most simple and elegant principle in Mathematics is the principle of duality. It arose from its applications in geometry, and it applies to classes of problems. In optimization theory, the dual of "minimization" is " maximization". Here, duality means reciprocity. Steinitz (1916) suggested the word correlation. This duality was probably, at least implicitly, known in Fermat's times for the problem of maximis and minimis of an unconstrained function. It sufficed to change the sign of the function. But, from all-important XIX$^{\text{th}}$ century contributors to the concept of duality, we retained two names. Joseph Diaz Gergonne (1771–1859) because he discovered the fundamental meaning of duality and Von Staudt (1798–1867) with his Geometrie der Lage, 1847, because of the strong impact it had in Germany. For Gergonne, the principle of duality was sketched by Euler:

> Except for some theorems, such as for instance Euler's, in the statement of which the number of faces and the number of vertices enter in the same way, there is no theorem of this kind which should not inevitably correspond to another, which can be deduced from it by merely exchanging the words faces and vertices with one another.

For Gergonne, the duality in Geometry indicates a double aspect in a proposition: faces and vertices in a polyhedron, or if on a given straight line, we can conceive an infinity of points, we return the proposition, on a given point, we can conceive an infinity of straight lines. If from two points, we can draw a straight line, the intersection of two straight lines is a point. If, one of the most famous examples of duality in geometry came from Desargues' theorem, Kepler, in 1619, talked about the "sexual" properties of platonic solids. The application of duality, this metathesis to LP problems will be more complex. Because duality was a hot topic in topology, duality was certainly familiar to topologists such as A. W. Tucker. Here, the duality applies to the problem of minimization (maximization) with the inequalities constraints. The key to duality will come from the old lagrangian technique in transforming a constrained problem into an unconstrained problem, and from the calculus of variations. In variational problems, duality relations are based on the Legendre transform.

Duality in LP will be introduced by von Neumann, Gale, Kuhn and Tucker, with full credit to John von Neumann. John von Neumann recognized the min-max problem. The beauty came from the bilinear symmetry between the variables and the lagrangian multipliers. And Kuhn (1976) said with humour "this duality, although it was discovered and explored with surprise and delight in the early days of linear programming, has ancient and honourable ancestors in pure and applied mathematics".

Indeed, the recent history of linear programming and its links with Operational Research, are well known (Dantzig 1963, Kuhn 1976, Fenchel 1983, Kjeldsen 1999). In particular, G. Dantzig, the leading person on Linear Programming (LP) in the USA published several testimonies. Duality was implicit in the 1873 Gordan's article. His article was rediscovered several times, and we wanted to quote these selected following reflections from Dines, in 1936, who came also very close to the discovery of duality:

> The theorems which we have just obtained may perhaps be described in a general way as *matrix free* theorems concerning adjoint systems of linear conditions. Two adjoint systems (from the transpose matrix) arose from the same matrix. The properties of the matrix determine the nature of the solution of each system. But once the characterization has been established, the matrix may be eliminated from consideration, and there results a relationship between the natures of the solutions of the adjoint systems.

However, duality in LP is more complex than just taking the adjoint of a matrix; it is obtained in a finite number of steps: transpose the coefficients of the matrix, interchanging the role of the constant terms and the coefficients of the objective function, changing the direction of inequality, and maximizing instead of minimizing, with anti-symmetry processes. Moreover, duality helped the understanding of LP problems. It brought the attention on the existence and uniqueness of solutions, on one hand the algebraic problem, and for example, the following table shows the correspondence between the solutions (or the absence of solution) of primal and dual problems; and on the other hand the problem of algebraic geometry with feasibility regions and Steinitz convex cones. The beauty of the geometrical representations of systems of inequalities, convex cones and their duals will appear in David Gale's book The Theory of Linear Economic Systems (1960).

Table 1 – Correspondence of solutions between the primal and the dual problem (from Papadimitriou and Steiglitz, 1982)

Primal \ Dual	Finite optimum	Unbounded	Infeasible
Finite optimum	(1)	X	X
Unbounded	X	X	(3)
Infeasible	X	(3)	(2)

At this step, important acquisitions came, not from the problem of duality, but from its applications of linear algebra with systems of inequalities, and indeed also from the modeling, from the fact that a simple linear model associated to a system of linear inequalities could have so many important applications in so many different fields such as military, economy, and industrial applications.

6 The art of computation

We saw that mathematicians found their aspirations in the calculus of variations, geometrical inequalities and algebraic geometry, linear algebra, the theory of games, duality in topology, network theory and the practical applications. The success of LP had a direct and encouraging influence on non-linear programming, with for example the Kuhn Tucker conditions, in 1951. Also, The linear hypothesis has always attracted Statisticians. Linear models became increasingly important as we considered more and more complicated experimental designs, because the linear links between variables corresponded to a principle of uncertainty. We also can find a similar cognition in LP. In both cases, we also have to solve systems of linear equations, and LP made an extensive use of the gaussian elimination algorithm, developed by Gauss in 1823–1826 for the least squares problem. The term "robust" was suggested by a statistician, G. E. P. Box in 1953, with the meaning being insensitive to small departures from the idealized assumptions (sensitivity to data). For example, we found a similar approach to the addition or deletion of variables in multiple-linear regression and LP. In parallel, the digital computer has provoked the birth of computer arithmetic and the art of scientific programming. Again, key articles came in 1947 from J. von Neumann and H. H. Goldstine, and from A. Turing. G. Dantzig's algorithm, the simplex method dated from 1951. The simplex method follows a sequence of vertices. It is a combinatorial approach. With no convergence criteria, it produces the answer in a finite time, but the number of steps (unlike the gaussian elimination) is not completely fixed, because we cannot tell in advance how many vertices the method will try. It does not possess the property of polynomial complexity.

References

- Dantzig, G. B., 1963. *Linear programming and extensions*, Princeton : USA, Princeton University Press.
- De la Vallée Poussin, Ch., 1911, "Sur la méthode de l'approximation minimum", *Annales de la société scientifique de Bruxelles*, vol. 35, pp. 1–16.
- Eggleston, H. G., 1958, *Convexity*, Cambridge at the University Press.
- Fenchel, W., 1953, *Convex cones, sets, and functions*, Princeton University.
- Fenchel, W., 1983, "Convexity through the ages", in *Convexity and Its Applications*, Basel : Birkhaüser Verlag, pp. 120–129.
- Fourier, J. B., 1798, "Mémoire sur la statique", *J. de l'école Polytechnique*, Ve cahier, p. 20, 1798.
- Fourier, J. B., "Solution d'une question particulière du calcul des inégalités, original 1826 paper with an abstract of an 1824 paper reprinted in *Œuvres* de Fourier", Tome II, Gauthier-Villars, Imprimeurs-Libraires, Paris : M DCCC XC, p. 317–328.
- Gale, D., 1960, *The theory of Linear Economic Models*, New York : McGraw-Hill.
- Kantorovich, L. V., 1939, "Mathematical methods in the Organization and Planning of Production, translated", in *Management Science* 6, pp. 366–422, 1960.
- Kjeldsen, T. H., 1999, "The Origin of Nonlinear Programming", *Proceedings of the CSHPM.*
- Kuhn, H. W., 1976, "Nonlinear Programming: A Historical View", "SIAM – AMS" *proceedings*, Vol. IX, pp. 1–26.

– Minkowski, H., 1910, *Geometrie der zahlen*, Leipzig : 1896, second edition 1910.

– Papadimitriou, C. H., Steiglitz, K., 1982, *Combinatorial Optimization*, Prentice-Hall.

– Steinitz, E., 1916, “Bedingt Konvergente reihen und Konvexe Systeme III”, *J. für die Reine und Angewandte Mathematik*, pp. 1–52.