Mathematical Methods of Economics

Joel Franklin
California Institute of Technology, Pasadena, California 91125

The American Mathematical Monthly, April 1983, Volume 90, Number 4, pp. 229–244.

When Dr. Golomb and Dr. Bergquist asked me to give a talk on *economics*, my first impulse was to try to get out of it.

"Sol," I said, "I'm not an economist. You know that."

"I know," said Golomb.

"If you want an economist, I can get you one," I said. "I know some excellent economists."

"No," he said, "we want a mathematician to talk about the subject to other mathematicians from their own point of view."

That made sense, and I hit on this idea: I won't try to tell you what mathematics has done for economics. Instead, I'll do the reverse: I'll tell you some things economics has done for mathematics. I'll describe some mathematical discoveries that were motivated by problems in economics, and I'll suggest to you that some of the new mathematical methods of economics might come into your own teaching and research.

One of these methods is called *linear programming.* I learned about it in 1958. I had just come to Caltech as a junior faculty member associated with the computing center. The director and I made a cross-country trip to survey the most important industrial uses of computers. In New York, we visited the Mobil Oil Company, which had just put in a multi-million-dollar computer system. We found out that Mobil had paid off this huge investment in *two weeks* by doing linear programming.

Back at Caltech, Professor Alan Sweezy in economics and Professors Bill Corcoran and Neil Pings in chemical engineering urged me to teach a course in linear programming. When I told them I didn't know linear programming, they said: Fine, Joel, *learn* it. Seeing they meant business, I did study the subject and give the course. The students loved it, and so did I. Perhaps you will have a similar experience.

One surprising thing I found was this: The mathematics was delightful. I knew it was useful, but I hadn't expected it to be beautiful. I was surprised to find that linear programming wasn't just business mathematics or engineering mathematics; it was the general mathematics of linear inequalities. Later I found this mathematics coming into some of my own special fields of research (statistics, numerical analysis, ill-posed problems). Here again, you may have a similar experience.

The author is Professor of Applied Mathematics at the California Institute of Technology. At Stanford University, in 1953, he received his Ph.D. in pure mathematics. His published research pertains to numerical analysis, to ill-posed problems, to stochastic processes, and to mathematical problems in engineering, in crystallography, in geophysics, and in cell biology. At Caltech he has twice received the Associated Students Award for Excellence in Teaching. He has written two textbooks for undergraduates in mathematics: *Matrix Theory* and *Methods of Mathematical Economics*.

*Invited address to the Mathematical Association of America and to the Society for Industrial and Applied Mathematics, November 14, 1981, Santa Barbara, California.

Linear programming is one of the many mathematical methods of economics. Here are a few others: quadratic programming, geometric programming, general nonlinear programming; fixed-point theorems—especially the Kakutani theorem; calculus of variations, control theory, dynamics programming; theory of convex sets—especially convex cones; probability, statistics, stochastic processes; finite structures (graph theory, lattice theory); matrix theory; calculus, ordinary differential equations; and special topics like game theory and Arrow's theory of rational preference orderings.

Plato said mathematics is the essence of reality; Willard Gibbs said mathematics is the language of science. If they are right, we shouldn't be surprised to find uses for any branch of mathematics in any science. Every branch of mathematics may have some use in the science of economics. Here are two bizarre examples:

Have you heard of *nonstandard analysis*? I've heard of it, but know next to nothing about it. Nevertheless, on November 10, 1981, I heard Yale Economics Professor Donald J. Brown give a colloquium on the *nonstandard analysis of hyper-finite economies* (see [**4**] and [**20**]).

You have heard of Bourbaki; so have I. I always thought that stuff would never be good for anything. Nevertheless, Bourbaki *ultrafilters* appear in a paper in the *Journal of Economic Theory* [**17**]. The authors, A. Kirman and D. Sondermann, use ultrafilters to generalize Kenneth Arrow's fundamental theorem of welfare economics [**1**].

Mathematics appears in all parts of economics, especially in *mathematical economics* and in *econometrics*. Mathematical economics is like mathematical physics: it is theoretical, nonempirical, sometimes speculative. For instance, Alfred Marshall hypothesized the *existence* of certain curves (supply and demand schedules) whose intersections determine commodity prices. Very pretty, but he didn't show how to measure or predict *numerical* values for specific supply-demand schedules.

In general, *measurement* and *prediction* belong to *econometrics*. As you would expect, econometrics uses a lot of mathematical statistics, probability theory, and numerical analysis. A Nobel prize was given in 1980 to Lawrence Klein for his work in building econometric models.

In 1969 the first Nobel prize in economics was given to Ragnar Frisch and Jan Tinbergen "for having developed and applied dynamic models for the analysis of economic processes"; in other words, the prize was given for mathematics applied to economics. Later, I'll show you a list of all the Nobel prizes in economics, and you'll see that at least 7 of the 12 prizes given from 1969 through 1981 were given for work that could be called applied mathematics. In fact, in 1975 a Nobel prize in economics was given to Leonid Kantorovich, who *is* a mathematician.

In 1969 a spokesman for the Nobel foundation welcomed the new prize subject, economics, as "the oldest of the arts, the youngest of the sciences." It might be fair to say that economics became a science when it started making significant use of mathematics. When was that? I'd say the nineteenth century.

In 1817 the stockbroker David Ricardo proved a *theorem* that establishes an astounding principle of international economics. Ricardo proved mathematically that *free trade is* (*under certain assumptions*) *advantageous to consumers in all nations.*

Alfred Marshall was another great nineteenth-century economist. Marshall started out to be a mathematician; he was First Wrangler in mathematics at Cambridge. Although his work is seldom explicitly mathematical, any mathematician reading it can sense its mathematical core. Marshall was a teacher of John Maynard Keynes, whose work contains plenty of explicit mathematics. But, at least to my taste, Marshall's work shows more mathematical insight.

As Gerard Debreu wrote in his *Theory of Value* [**7**], mathematical economics has become increasingly geometric and qualitative. If we want precise numerical information, we have to turn to *econometrics*. Whereas Marshall drew his supply-and-demand curves in a nonnumerical, qualitative way, the econometrician would have the hard problem of giving *numerical* values for these curves for specific commodities at specific times.

An example of econometrics appears in an article [**29**] by mathematician Jacob Schwartz. He used a Wharton econometric model for residential housing. You can see it in Fig. 1. There you see a typical awful equation of econometrics; please don't try to understand it. I just want you to see what is looks like. It predicts the rate of investment in residential housing as a function of various factors (the numerical subscripts refer to time lags). The coefficients (58.26, 0.0249, etc.) come from a numerical curve fit to data for 1948–1964; the model was published in 1967.

There is an old Chinese proverb: *It is always difficult to predict—especially the future*. For that reason econometrics is difficult. The Wharton model of 1967 "predicts" housing starts for 1948–1964—not for the future. In general, econometric models are not laws of nature like $f = ma$ or $E = mc^2$; they are empirical studies whose predictive value depends on the constancy of the underlying relationships.

1967 Wharton econometric model (for 1948–1964)

$$I_h = 58.26 + 0.0249Y - 45.52\left(\frac{p_h}{p_r}\right)_{-3} + 1.433(i_L - i_s)_{-3} + 0.0851(I_h^s)_{-1}$$

I_h = rate of investment (\$$10^9$) in residential housing per quarter (3 months)

Y = total disposable income

p_h = average housing price

p_r = average rental price

i_L = long-term interest rate

i_s = short-term interest rate

I_h^s = rate of housing starts

Negative subscripts denote time lags.

FIG. 1.

What Do Economists Think of Mathematics? That question has had different answers at different times. *Now* the answer would be overwhelmingly favorable, if not unanimous. But not so in the old days. Adam Smith published his great book *Wealth of Nations* in 1776. It is readable, fascinating, and important; but it contains almost no mathematics.

I told you the great nineteenth-century economist Alfred Marshall had been First Wrangler in mathematics at Cambridge. Later, he talked about the role mathematics played in his work:

> I had a growing feeling in the later years of my work at the subject that a good mathematical theorem dealing with economic hypotheses was very unlikely to be good economics: and I went more and more on the rules—(1) Use mathematics as a shorthand language, rather than as an engine of inquiry. (2) Keep to them till you have done. (3) Translate into English. (4) Then illustrate by examples that are important in real life. (5) Burn the mathematics. (6) If you can't succeed in 4, burn 3. This last I did often. —quoted in [**31**], p. 307.

So Marshall practiced mathematics as a secret vice; he was a closet mathematician. His most famous student was John Maynard Keynes. At Cambridge, Keynes took his degree in mathematics*. In 1920 Keynes published his *Treatise of Probability*. Keynes's great books on economics contain many equations. By the time of Lord Keynes mathematics was not a secret vice but a public virtue.

A living disciple of Keynes, Harvard Professor John Kenneth Galbraith, regards mathematics with skepticism. One of Galbraith's more entertaining books is called *Economics, Peace, and Laughter*. Commenting on the models of mathematical economics, he says this:

> Moreover, the models so constructed, though of no practical value, serve a useful academic function. The oldest problem in economic education is how to exclude the incompetent . . . . The requirement that there be an ability to master difficult models, including ones for which mathematical competence is required, is a highly useful screening device.

Not satisfied with this comment, Galbraith adds a dour footnote:

> There can be no question, however, that prolonged commitment to mathematical exercises in economics can be damaging. It leads to the atrophy of judgment and intuition . . . .

John Galbraith does not stand alone. He tells this story about Paul Samuelson, a superb applied mathematician and winner of the Nobel Prize for work in mathematical economics:

> Professor Samuelson, in his presidential address to the American Economic Association several years ago, noted that the three previous presidential addresses had been devoted to a denunciation of mathematical economics and that the most trenchant had encouraged the audience to standing applause.

Well! And skepticism about mathematics is not confined to this continent. Galbraith says:

> Once when I was in Russia on a visit to Soviet economists, I spent a long afternoon attending a discussion on the use of mathematical models in plan formation. At the conclusion an elderly scholar, who had also found it very heavy going, asked me rather wistfully if I didn't think there was still a "certain place" for the old-fashioned Marxian formulation of the labor theory of value.

---

*While studying for the Tripos, Keynes wrote to his friend B. W. Swithinbank on 18 April 1905: "I am soddening my brain, destroying my intellect, souring my disposition in a panic-stricken attempt to acquire the rudiments of the Mathematics." See R. F. Harrod [**13**], p. 130.

The old Russian scholar must have sighed when a Nobel prize in economics was given to Leonid Kantorovich, a mathematician. Kantorovich got the prize for developing the mathematical theory of linear programming and for applying it to the economic problem of optimum allocation of resources. He would have gone a lot farther with linear programming if he hadn't run into trouble from the orthodox Marxians, who objected to the use of the idea of prices. Dantzig tells the story in his book [**6**], p. 23.

Among the Nobel Laureates in economics, some, like Kantorovich, solved problems in economics by inventing new mathematics; others made much use of known mathematics. Look at the list of Nobel prizes in economics, Fig. 2. I've put asterisks by seven of the twelve prize years to indicate work that is heavily mathematical.

Nobel Prizes in Economics

1969* Frisch, Ragnar and Tinbergen, Jan—"for having developed and applied dynamic models for the analysis of economic processes."

1970* Samuelson, Paul—"for the scientific work through which he has developed static and dynamic economic theory and actively contributed to raising the level of analysis in economic science."

1971 Kuznets, Simon—"for his empirically founded interpretation of economic growth which has led to new and deepened insight into the economic and social structure and process of development."

1972* Hicks, Sir John R. and Arrow, Kenneth J.—"for their pioneering contributions to general economic equilibrium theory and welfare theory."

1973 Leontief, Wassily—"for the development of the input-output method and for its application to important economic problems."

1974 Myrdal, Gunnar and Von Hayek, Friedrich August—"for their pioneering work in the theory of money and economic fluctuations and for their penetrating analysis of the interdependence of economic, social and institutional phenomena."

1975* Kantorovich, Leonid and Koopmans, Tjalling—"for their contributions to the theory of optimum allocation of resources."

1976* Freidman, Milton—"for his achievements in the fields of consumption analysis, monetary history and theory and for his demonstration of the complexity of stabilization policy."

1977 Ohlin, Bertil and Meade, James—"for their pathbreaking contributions to the theory of international trade and international capital movements."

1978 Simon, Herbert A.—"for his pioneering research into the decision-making process within economic organizations."

1979 Lewis, Arthur and Shultz, Theodore—for studies of human capital.

1980* Klein, Lawrence—for computer models designed to forecast economic changes.

1981* Tobin, James—for mathematical models of investment decisions.

* Asterisks indicate very mathematical work.

FIG. 2.

Seven out of twelve Nobel prizes—not a bad score for mathematics. And some of this mathematics has freshness and charm. For example, let me show you a theorem that won a Nobel prize: the Possibility Theorem of Kenneth Arrow.

In 1957 Kenneth Arrow published a little book called *Social Choice and Individual Values*. He was thinking about a problem of welfare economics: Confronted by numerous conflicting special interests, how should the government make decisions?

Use old-fashioned majority rule, you say. That's the democratic way isn't it? That's the *rational* way.

Let's see. Suppose we have 3 *alternatives*: vanilla (V), chocolate (C), and strawberry (S). And suppose we have 9 voters, each with his own *individual values*. For example, one individual may like vanilla better than chocolate $(V > C)$, and he may like chocolate better than strawberry $(C > S)$; then, by the way, he must like vanilla better than strawberry $(V > S)$ if his individual values are *rational*. Another individual may prefer strawberry to vanilla $(S > V)$, vanilla to chocolate $(V > C)$, and *therefore* strawberry to chocolate $(S > C)$. And so on.

If all of our nine voters have definite flavor preferences, the voters constitute 6 special-interest groups, corresponding to the six ways of ranking 3 flavors. For example, we might have the following tabulation:

Individual values	Number of individuals
$V > C > S$	2
$S > V > C$	2
$C > S > V$	2
$V > S > C$	1
$C > V > S$	1
$S > C > V$	1

Now comes the general election. Here are the results:

$V > C$ by a majority of 5 to 4

$C > S$ by a majority of 5 to 4

and—what's this?

$S > V$ by a majority of 5 to 4.

But that's crazy: $V > C$ and $C > S$ *should* imply $V > S$, not $S > V$. (This is an example of *Concordet's paradox*.)

No wonder Congress is confused. You see the problem. So did Arrow, and he wondered if there was any way out.

There *is* one way out: Hitler's way. Pick one individual, call him *der Fuhrer*, and do what he says. Then all the government's preferences can be nice and transitive, and too bad for you if you don't like it.

Is there any rational way to make social choices besides dictatorship? To this basic question of welfare economics, Kenneth Arrow gave an astonishing answer: *No*.

ARROW'S THEOREM. *Suppose we have a function that makes rational (transitive) social choices as a function of rational individual values that rank (by preference or indifference) three or more alternatives. Assume that the social-choice function has two properties:*

(i) *If all individuals prefer alternative a to alternative b, then society shall prefer a to b.*

(ii) *The social choice between any two alternatives a and b shall depend only on the individual values between a and b (and should not depend on any third alternative c).*

Then Arrow's theorem says there exists a dictator—a single individual whose preferences become social choices.

In a minute I'll write this theorem symbolically, in terms of matrices. But first I want to explain the two assumptions. The first is a principle of *unanimity*: If everyone prefers vanilla to chocolate, so should society. The second is a principle of *relevance*: Society's choice between vanilla and chocolate should depend on how people feel about vanilla and chocolate, not on how they feel about strawberry.

If you wish, you can write Arrow's theorem in terms of matrices. Let $a_{ij} = 1$ if i is preferred to j; let $a_{ij} = -1$ if j is preferred to i; let $a_{ij} = 0$ if neither is preferred to the other. If there are m alternatives (flavors), then the numbers a_{ij} constitute an $m \times m$ skew-symmetric matrix, A. In a rational preference ordering, if i is preferred to j, and if j is preferred to k, then i must be preferred to k. For the matrix A this says: If $a_{ij} = 1$ and $a_{jk} = 1$, then $a_{ik} = 1$. We shall also require $a_{ik} = 1$ if $a_{ij} = 0$ and $a_{jk} = 1$ *or* if $a_{ij} = 1$ and $a_{jk} = 0$. If this is so, then we'll call A a *rational preference matrix*.

EXAMPLE. Suppose we prefer flavor 3 to flavor 1 and flavor 2, which we like equally. Then this is our rational preference matrix:

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

EXAMPLE. Suppose we prefer flavor 1 to flavor 2, flavor 2 to flavor 3, *and* flavor 3 to flavor 1. That is irrational, and so the preference matrix is *irrational*:

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Look: $a_{12} = 1$ and $a_{23} = 1$, *but* $a_{13} \neq 1$.

Individual values and social choice: Suppose there are n individuals and m alternatives. The individual values are expressed by n rational preference matrices$A(1),\ldots, A(n)$. A social choice is a rational preference matrix A. We're looking for a *function* F mapping P_m^n into P_m, where P_m is the set of $m \times m$ rational preference matrices and P_m^n is the n-fold Cartesian product:

$$A = F(A(1),\ldots, A(n)).$$

EXAMPLE. For majority rule, the function F is defined as follows:

$$a_{ij} = \text{sign}[a_{ij}(1) + \cdot \cdot \cdot + a_{ij}(n)] \quad (i, j = 1, \ldots, m).$$

If $m > 2$, majority rule may give irrational social choices, as we saw in the example of vanilla, chocolate, and strawberry. So this F takes values outside P_m; but this F does satisfy the assumption of *unanimity* and *relevance*:

(1) $a_{ij} = 1$ if $a_{ij}(k) = 1 \forall k = 1, \ldots, n$

(2) $a_{ij} =$ a function of $a_{ij}(1), \ldots, a_{ij}(n)$.

Arrow's theorem now takes this form: *Let F be a function mapping* P_m^n *into* P_m. *Suppose* $m > 2$, *and suppose the function F satisfies equations* (1) *and* (2). *Then there exists an integer d such that* $a_{ij} = 1$ *if* $a_{ij}(d) = 1$. (*The integer d depends on F but not on the matrices* $A(1), \ldots, A(n)$.)

By the way, there are no restrictions on the number of individuals, n. In marriage, $n = 2$. Then Arrow's theorem says: Either the husband or the wife must be a dictator, or there must be irrational choices. Experience seems to bear this out.

Arrow's theorem talks about *rational* (transitive) preference orderings. This raises a question in combinatoric analysis: How may rational preferences orderings of m alternatives are there? The answer has appeared in [**12**]. For large m the number of rational preference orderings behaves like $(1/2)m!(\log 2)^{-m-1}$.

The mathematics of Arrow's theorem is very different from mathematics like linear programming. Here we have a rather ordinary looking problem:

For $i = 1, \ldots, m$ and $j = 1, \ldots, n$ we are given the real numbers a_{ij}, b_i, c_j. We wish to find numbers $x_j \geqslant 0$ such that

$$\sum_{j=1}^{n} c_j x_j = \text{minimum}$$

over all solutions of the linear equations

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \quad (i = 1, \ldots, m).$$

That is the canonical form of linear programming. In terms of matrices and vectors, it looks like this:

$$Ax = b, \quad x \geqslant 0, c^T x = \text{min}.$$

The problem is interesting only if the linear system $Ax = b$ has more than one solution, so we usually suppose rank $A = m < n$. Then the crucial assumption is the sign constraint $x \geqslant 0$ (all components of x must be nonnegative).

Kantorovich in Russa and Dantzig in the United States independently developed linear programming to solve economic logistical problems. The history of their work appears in Dantzig's book [**6**].

The most famous early problem of linear programming, the *diet problem*, first appeared in the *Journal of Farm Economics* [**33**]. The problem is to design a nutritionally adequate diet at minimum cost. The author, George Stigler, won the 1982 Nobel Prize in Economics.

Suppose a_{ij} is the amount of nutrient i in one unit of food j. (For instance, a_{37} might be the amount of vitamin B_1 in one gram of wheat bread.) Let b_i be the minimum daily requirement of nutrient i, and let c_j be the cost of one unit of food j. Let x_j be the amount of food j in a daily diet. Then we require

$$\sum_{j=1}^{n} a_{ij}x_j \geqslant b_i (i = 1,\ldots, m), \quad x_j \geqslant 0 \quad \sum_{j=1}^{n} c_j x_j = \text{minimum.}$$

This is a linear program in *standard* form. To put it in canonical form, we must replace the m linear inequalities by equations. We do that by introducing m new unknowns $z_i \geqslant 0$:

$$\sum_{j=1}^{n} a_{ij}x_j - z_i = b_i.$$

The problem is now easy to solve by Dantzig's *simplex method*.

Linear programming has many uses in industry and banking. In 1981, a good popular article [**2**] appeared in *Scientific American*; I recommend its example on beer. An introduction to the use of linear programming for the optimization of bank investment portfolios appeared in the *Monthly Review* of the Federal Reserve Bank of Richmond (see [**3**] and [**10**], p. 3). Banks and oil companies make a lot of money with linear programming.

But you and I are mathematicians; money means nothing to us. So let us speak of something more important—let's talk about *Chebyshev approximation*.

Suppose we are given a system of real linear equations, $Ax = b$, and suppose the system has no solution x. Typically, this occurs when we have more equations than unknowns. If we have m equations in n unknowns, the error in equation i is a function of the vector x:

$$e_i = \sum_{j=1}^{n} a_{ij}x_j - b_i \quad (i = 1,\ldots, m).$$

The problem of Chebyshev approximation is to find a vector x that minimizes the maximum absolute error:

$$\operatorname*{Minimize}_{x} \left(\max_{i} |e_i|\right).$$

That is a beautiful and important problem of approximation theory. Many things were known about Chebyshev approximation before 1959, but no one knew a good way to *do* it. Then Edward Stiefel discovered how to do it by linear programming (see [**32**] and [**10**], p. 8). Here's how:

Define a new unknown: $x_0 = \max|e_i|$ for $i = 1,\ldots, m$. Then we shall have the uniform error bracket

$$-x_0 \leqslant \sum_{j=1}^{n} a_{ij}x_j - b_i \leqslant x_0 \quad (i = 1,\ldots, m).$$

The problem of Chebyshev is to choose $x_0,\ldots, x_n$ so as to minimize the maximum absolute error: Minimize x_0.

That's all there is to it—a finite number of linear inequalities in a finite number of unknowns, with a linear form to be minimized. That is a linear program in general form.

It's trivial to restate it in canonical form, and it's routine to solve it numerically by the simplex method.

The simplex method is perhaps the most important numerical method invented in the twentieth century. Experience with enormous industrial problems shows that the simplex method works *fast*. In problems with m equations in n unknowns, the computation time seems to be proportional to n.

Why does the simplex method usually work so fast? No one knows, and this is one of the great unsolved problems of numerical analysis. At first glance, the computation time would seem to be proportional to the binomial coefficient $\binom{n}{m}$, which is the possible number of basic solutions of $Ax = b$. For $m \sim n/2$, the binomial coefficient is almost as big as 2^n, and this suggests the computing time could grow exponentially with n. Indeed, Victor Klee and George Minty [**18**] have constructed pathological cases for which that happens. But it never seems to happen in practice.

A Russian mathematician named Khachian got around this problem by analyzing a quite different algorithm [**16**]. Khachian proved that his algorithm has computing time bounded by a constant, K, times n^6—which becomes smaller than 2^n. Khachian's proof is a triumph of theoretical computer science. But Khachian's algorithm, in its present form, has little practical value: the constant K is enormous and so is the computing time.

You can become famous by doing one of these two things: (1) show why the simplex method usually works as well as it does; (2) show how Khachian's method can be made to work *better* than the simplex method in practice. [A persistent rumor says Stephen Smale *has* done (1).]

Linear programming is important because it is the general mathematics of finite systems of linear inequalities. Linear programming is more general than real linear algebra, for this reason:

Any real linear equation $\sum a_i x_i = b$ can be restated as a pair of linear inequalities:

$$\sum a_i x_i \leqslant b \text{ and } \sum a_i x_i \geqslant b.$$

But the converse is false: You can't restate a linear inequality as a finite number of linear equations.

No mathematician doubts the importance of linear algebra. So linear programming must also be important, and perhaps you will agree that linear programming should be part of the basic undergraduate mathematics curriculum. Why should mathematics students have to pick up their linear programming from economists and chemical engineers and people like that? They should learn it from *us*, and they should learn it right.

Marshall Hall has a section on linear programming in his book *Combinatoric Analysis*. There's nothing odd about that; linear programming has many applications to combinatorics. For instance, look at this problem:

We are given an $n \times n$ matrix of real numbers a_{ij}. We seek a permutation $j_i, \ldots, j_n$ that maximizes the sum

$$s = a_{1j_1} + a_{2j_2} + \cdots + a_{nj_n}.$$

This problem is called the *optimal-assignment* problem.

EXAMPLE. Suppose we're given the matrix

$$\begin{pmatrix} 7 & 2 & 6 \\ 3 & 9 & 1 \\ 8 & 4 & 5 \end{pmatrix}.$$

The sum s has six possible values. The largest is

$$\max s = a_{13} + a_{22} + a_{31} = 6 + 9 + 8 = 23,$$

achieved for the permutation $(j_1, j_2, j_3) = (3, 2, 1)$.

In general, we could solve the problem by calculating all the $n!$ possible values for s, but that takes too long if n is large. A much faster algorithm is given by linear programming.

We define the unknowns x_{ij} as 1 if $j = j_i$, or 0 if $j \neq j_i$. Thus, x_{ij} will tell us which component to pick from each row. For the preceding numerical example, we would have

$$(x_{ij}) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

In general, the integer unknowns x_{ij} must satisfy the constraints

$$x_{i1} + \cdots + x_{in} = 1 \quad (i = 1, \ldots, n)$$

$$x_{ij} + \cdots + x_{nj} = 1 \quad (j = 1, \ldots, n)$$

$$x_{ij} \geqslant 0 \quad (i, j = 1, \ldots, n).$$

Then we wish to maximize a linear form:

$$s = \sum_{i,j} a_{ij} x_{ij} = \text{maximum}.$$

This is a problem in linear programming. H. W. Kuhn [**19**] has shown that it can be solved in $O(n^3)$ steps.

You are right if you object that linear programming provides the optimal *real* solution x_{ij}, and these numbers might not be integers (we need all $x_{ij} = 0$ or 1). But for the optimal-assignment problem the optimal solution over the integers x_{ij} is also optimal over the real numbers x_{ij}. That's not obvious, but it's easy to prove. In general, however, linear programming over the integers is difficult. The optimal solution over integers is usually *not* optimal over real numbers.

So much for combinatorics. Now let's look at geometry. I'd like to show you how *quadratic* programming solves a problem stated in 1857 by J. J. Sylvester [**34**]: "It is required to find the least circle which shall contain a given set of points in the plane."

Suppose the given points are $\mathbf{a}_1, \ldots, \mathbf{a}_n$. We're looking for a circle with the unknown center $\mathbf{x}$ and radius ρ. The given points are required to lie inside the circle:

$$\|\mathbf{a}_i - \mathbf{x}\|^2 \leqslant \rho^2 \quad (i = 1, \ldots, m).$$

Then we want to choose $\mathbf{x}$ and ρ so as to minimize ρ.

We can replace the m quadratic inequalities by *linear* inequalities as follows. Introduce the unknown

$$x_0 = \frac{1}{2}(\rho^2 - \|\mathbf{x}\|^2).$$

Then the m inequalities become

$$x_0 + \mathbf{a}_i \cdot \mathbf{x} \geqslant b_i \quad (i = 1,\ldots, m),$$

where $b_i = \frac{1}{2}\|\mathbf{a}_i\|^2$. Then we want to minimize ρ^2:

$$2x_0 + \|\mathbf{x}\|^2 = \text{minimum}.$$

Sylvester's problem now has this form: First we require m linear inequalities:

$$x_0 + a_{i1}x_1 + a_{i2}x_2 \geqslant b_i \quad (i = 1,\ldots, m).$$

Then we want

$$2x_0 + x_1^2 + x_2^2 = \text{minimum},$$

in which the quadratic terms constitute a positive definite form. This is a routine problem of quadratic programming. It can be solved numerically by an ingenious variant of the simplex method. This algorithm was discovered by a mathematician, Philip Wolfe, but it was published in an *economics* journal, *Econometrica* [**36**].

Why in an economics journal? Because Wolfe's paper extended the work of some economists who were interested in the use of quadratic programming to make optimal investment decisions. Wolfe's mathematical discovery solved a problem in economics.

The theoretical basis of linear and nonlinear programming was published in 1902 by a mathematician named Julius Farkas. He gave a long, cumbersome proof of the following proposition, which you might call *the alternative of linear inequalities* (generalizing the Fredholm alternative of linear equations):

THE FARKAS THEOREM. *Let A be a given $m \times n$ real matrix, and let b be a given vector with m real components. Then one, and only one, of the following alternatives is true:*

(i) *the system $Ax = b$ has a solution $x \geqslant 0$ (all components $\geqslant 0$);*

(ii) *the system of inequalities $y^TA \geqslant 0$ has a solution y satisfying $y^Tb < 0$.*

Indeed, *both* alternatives can't be true, for then we could deduce

$$0 \leqslant (y^TA)x = y^T(Ax) = y^Tb < 0.$$

That's easy; the hard part is to show that *one* of the alternatives must be true. A modern straightforward proof of the Farkas theorem relies on the separating-plane theorem for convex sets (see, e.g., [**10**], p. 56).

The Farkas alternative has many uses outside mathematical economics. I hope to convince you that every mathematician should know the Farkas theorem and should know how to use it. For example, let me show how to use the Farkas theorem to prove the fundamental theorem of finite Markov processes.

THEOREM (Markov). *Suppose* $p_{ij} \geqslant 0$, *and suppose*

$$\sum_{i=1}^{n} p_{ij} = 1 \quad (j = 1, \ldots, n).$$

Then there exist numbers $x_j \geqslant 0$ *satisfying*

$$\sum_{j=1}^{n} p_{ij} x_j = x_i \quad (i = 1, \ldots, n)$$

$$\sum_{j=1}^{n} x_j = 1.$$

The proof of a special case of this theorem occupies several pages in Feller's book on probability ([**8**], pp. 428–432). The general case is usually proved by using the Perron-Frobenius maximum principle for positive matrices or by using the Brouwer fixed-point theorem. Instead, we can give an elementary proof using the Farkas theorem ([**10**], p. 58):

First, we state Markov's assertion as one Farkas alternative:

(i) There exists a vector $x \geqslant 0$ satisfying the $n + 1$ linear equations

$$\sum_{j=1}^{n} (p_{ij} - \delta_{ij}) x_j = 0 \quad (i = 1, \ldots, n)$$

$$\sum_{j=1}^{n} x_j = 1,$$

where δ_{ij} is the Kronecker delta.

Second, we state the other Farkas alternative:

(ii) There exist numbers $y_1, \ldots, y_n, y_{n+1}$ satisfying the inequalities

$$\sum_{i=1}^{n} y_i (p_{ij} - \delta_{ij}) + y_{n+1} \geqslant 0 \quad (j = 1, \ldots, n)$$

$$y_{n+1} < 0.$$

Alternative (ii) implies the strict inequalities

$$\sum_{i=1}^{n} y_i p_{ij} > y_j \text{ for all } j.$$

But

$$\max y_i \geqslant \sum_{i=1}^{n} y_i p_{ij}$$

because we assumed $p_{ij} \geqslant 0$ and $\Sigma_i p_{ij} = 1$, so we find

$$\max y_i > y_j \text{ for all } j.$$

That is impossible, so alternative (ii) is false.

Now Farkas tells us that alternative (i) is true: Markov's theorem is proved. That was easy, wasn't it?

Now let me tell you about *the theory of games and economic behavior*. A book with that title was published in 1944 by the mathematician John von Neumann and the economist Oskar Morgenstern [**25**]. Economists consider this book an epoch-making contribution to economics.

Fine, you say, but *what has it done for mathematics*?

This book, along with von Neumann's earlier work [**24**] on game theory, has given us some stimulating problems and some important results. For example, look at this theorem on matrices:

THEOREM (VON NEUMANN). *Let A be a real* $m \times n$ *matrix. Let vectors x and y range over the sets*

$$\sum_{i=1}^{m} x_i = 1, \quad x_i \geqslant 0; \quad \sum_{j=1}^{n} y_j = 1, \quad y_j \geqslant 0.$$

Then

$$\min_y \max_x x^T A y = \max_x \min_y x^T A y.$$

This theorem is no platitude. As a rule, mixed extrema are *not* equal, as the following example shows. Suppose x and y range over the sets $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. Then

$$\min_y \max_x (x - y)^2 = \frac{1}{4},$$

but

$$\max_x \min_y (x - y)^2 = 0.$$

Von Neumann's *minimax theorem* is the fundamental result in the theory of zero-sum two-person games. But that's not the point; the point is, it's good *mathematics*. Von Neumann proved the minimax theorem by using the Brouwer fixed-point theorem. His proof is nonelementary and nonconstructive. Later, the mathematician George Dantzig gave an elementary, constructive proof by using the dual simplex method of linear programming.

Following von Neumann, mathematical economists make much use of the fixed-point theorems. Their favorite seems to be the fixed-point theorem of Kakutani [**15**].

As a young mathematician at the Institute of Advanced Study, Shizuo Kakutani discovered a generalization of the Brouwer fixed-point theorem. Kakutani's work was motivated by problems in economic game theory. His theorem has great mathematical novelty. It speaks of point-to-set mappings:

THEOREM (Kakutani). *Let X be a closed, bounded, convex set in* $\mathbb{R}^n$. *For every point* x *in X, let* $F(\mathrm{x})$ *equal a nonempty convex subset of X. Assume that the graph*

$$\{\mathrm{x}, \mathrm{y} : \mathrm{y} \in F(\mathrm{x})\} \text{ is closed.}$$

Then some point in X satisfies $\mathrm{x}^* \in F(\mathrm{x}^*)$.

The image of each point x is a convex set $F(x) \subset X$. The theorem says some point x^* lies *in* its image $F(\mathrm{x}^*)$. Figure 3 illustrates this. Kakutani's theorem is novel because it talks about *set*-valued functions.

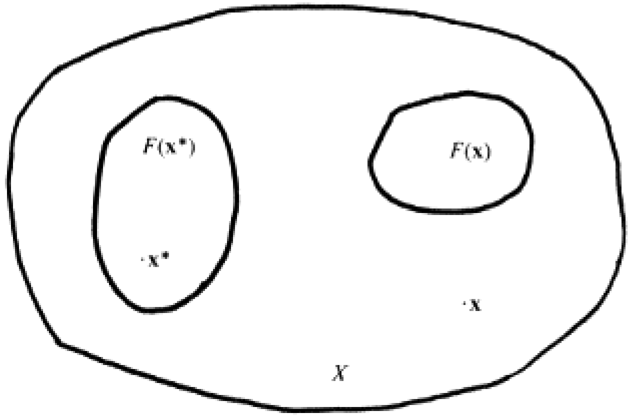

FIG. 3.

If every set $F(\mathrm{x})$ contains just one point, the closed-graph assumption is equivalent to the continuity of the function $F(\mathrm{x})$ and then Kakutani's theorem reduces to the Brouwer fixed-point theorem. Kakutani proved his theorem by using the Brouwer theorem.

A private survey indicates that 96% of all mathematicians can state the Brouwer fixed-point theorem, but only 5% can prove it. Among mathematical economists, 95% can state it, but only 2% can prove it (and these are all ex-topologists). This dangerous situation will soon be remedied. Within the last two years, John Milnor [**22**] and C. A. Rogers [**27**] have produced elementary proofs, using nothing more advanced than calculus. These proofs are so easy that *I* can understand them [**10**], and certainly *you* can.

While 96% of mathematicians can *state* the Brouwer fixed-point theorem, only 7% can state the Kakutani theorem. This situation is also dangerous, or, at least, wasteful. The Kakutani theorem has many potential applications outside economics; these applications should be made. Now that we can all understand the Brouwer theorem, we can also understand the Kakutani theorem, so nothing can stop us.

In the application of Kakutani's theorem to many-person game theory, the point x denotes a collection of mixed strategies and the set-valued function denotes the sets of optimal mixed strategies. The inclusion $x \in F(\mathrm{x})$ characterizes an equilibrium solution of the game. The Kakutani theorem is thus the perfect tool for proving J. F. Nash's fundamental theorem [**23**] on *n*-person games.

Professor H. F. Bohnenblust once told me something about research. He had supervised many successful Ph.D. thesis projects—and a few unsuccessful ones. He said this: The unsuccessful projects start with some famous old problem (prove the Riemann hypothesis) and *then* look for a method to solve it. The successful projects *start with some new method and then look for a problem*.

Let's take Bohnenblust's advice. Let's start with linear programming and look for a problem. Here's a good one: the *problem of moments* in probability theory.

Suppose we are given a collection of real-valued continuous functions $a_i(\mathrm{t})$ for $\mathrm{t} \in \mathbb{R}^p$. We are given a closed set $\Omega \subset \mathbb{R}^p$, and we're given a collection of real numbers b_i. The problem is to find a probability distribution function $x(t)$ satisfying the moment equations

$$\int_{\Omega} a_i(\mathrm{t})\, dx(\mathrm{t}) = b_i \text{ for all } i$$

where we require $dx(\mathrm{t}) \geq 0$ and

$$\int_{\Omega} dx(\mathrm{t}) = 1.$$

This problem has many applications in geophysics and in other sciences. It has an extensive mathematical theory (see, for instance, Shohat and Tamarkin [**30**]). So what is left for you and me to do here? Well, for one thing, we could devise a good numerical method. At least, that will please our colleagues in geophysics.

Suppose we're given a *finite* number of moments, which is the usual case in applications. And suppose we use some numerical scheme to approximate the integrals by finite sums. Then we get a finite set of linear equations in a finite set of unknowns:

$$\sum_{j=1}^{n} a_{ij} x_j = b_j \quad (i = 1, \ldots, m).$$

Now we're looking for the numbers $x_1, \ldots, x_n$; they will constitute a finite set of probabilities, satisfying

$$\sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0.$$

So we want to solve $m + 1$ linear equations in n unknowns $x_j \geq 0$. Ah! We recognize a problem in linear programming. For this we have an existence theorem, the Farkas theorem, and a numerical method, the simplex method.

The simplex method will tell us if no solution exists, or it will compute a solution x if solutions do exist. For $n > m + 1$ we can't expect the solution x to be unique. We are free to impose any minimum condition of the form

$$\sum_{j=1}^{n} c_j x_j = \text{minimum}.$$

We note that the original problem with a *finite* number of moments usually doesn't have a unique solution $x(\mathrm{t})$, so the freedom to impose an extra condition is physically natural and mathematically necessary.

Fine, you say. All right for some people but not for you. You are a pure mathematician, and numerical methods bore you. What you'd like is a little solid theory—something you can get your teeth into.

OK, I'm with you. Let's prove a great theorem together. Let's give a new, elementary proof of a famous theorem of F. Hausdorff [**14**]. The proof will use a method of mathematical economics, the Farkas theorem.

Hausdorff studied the moment problem

$$(3) \qquad \int_0^1 t^k\, dx(t) = b_k \quad (k = 0, 1, \ldots).$$

He asked this question: *Which infinite sequences* $\{b_k\}$ *are the moments of a probability distribution* $x(t)$ *on the interval* $0 \le t \le 1$? He called those sequences *moment sequences*.

Certainly $b_0 = 1$, since we require $\int dx(t) = 1$. Also we must have

$$\int_0^1 f(t)\, dx(t) \ge 0$$

for all continuous functions $f(t) \ge 0$. Setting $f(t) = t^j(1 - t)^k$, we get the necessary condition

$$\int_0^1 \sum_{\nu=0}^{k} (-)^\nu \binom{k}{\nu} t^{j+\nu}\, dx(t) \ge 0,$$

which says this about the moments:

$$\sum_{\nu=0}^{k} (-)^\nu \binom{k}{\nu} b_{j+\nu} \ge 0 \ (j, k \ge 0).$$

A sequence $\{b_i\}$ with this property is called *completely monotone*. If we define the difference operator Δ by $\Delta b_i = b_{i+1} - b_i$, the last formula says

$$(-)^k \Delta^k b_j \ge 0 \quad (j, k \ge 0).$$

Hausdorff's theorem says: *If* $b_0 = 1$, *the sequence* $b_0, b_1, b_2, \ldots$ *is a moment sequence if and only if it is completely monotone.*

We've already proved the *only if* part. To prove the *if* part, let's assume the sequence $\{b_i\}$ is completely monotone, with $b_0 = 1$. Now we must find a p.d.f (probability distribution function) $x(t)$ satisfying the moment equations (3).

Suppose we can solve the system of moment equations

$$\text{(i)} \quad \int_0^1 t^k\, dx_n(t) = b_k \quad (k = 0, \ldots, n)$$

for each *finite n*. Then the p.d.f.'s $x_n(t)$ have a subsequence that converges to a p.d.f. $x(t)$ at all points of continuity of the limit $x(t)$. Then $x(t)$ satisfies *all* the moment equations (3), and we're done.

So the required p.d.f. $x(t)$ exists unless some *finite* system (i) is unsolvable. But the system (i) is a finite linear system for an unknown $dx_n(t) \ge 0$. A simple extension of the Farkas theorem says this: *The system* (i) *is unsolvable for a* p.d.f. $x_n(t)$ *if and only if there exist numbers* $y_o, \ldots, y_n$ *satisfying*

$$\text{(ii)} \quad \sum_{k=0}^{n} y_k t^k \ge 0 \quad (0 \le t \le 1)$$

$$\sum_{k=0}^{n} y_k b_k < 0.$$

We must show that this is impossible.

Suppose (ii) is true. Define the polynomial $f(t) = \Sigma y_k t^k$. Then Taylor's theorem says

$$y_k = f^{(k)}(0)/k!$$

As a limit of difference quotients, this equals

$$y_k = \lim_{\varepsilon \to 0} \Delta_\varepsilon{}^k f(0)/(\varepsilon^k k!),$$

where $\Delta_\varepsilon f(t) = f(t + \varepsilon) - f(t)$. Setting $\varepsilon = 1/N$, we deduce

$$y_k = \lim_{N \to \infty} \binom{N}{k} \Delta_\varepsilon^k f(0).$$

The second part of (ii) says $\Sigma y_k b_k < 0$, and so for large N we must have

$$\sum_{k=0}^{n} \binom{N}{k} (\Delta_\varepsilon^k f(0)) \cdot b_k < 0.$$

The upper limit, n, may be replaced by a larger integer, N, since an nth degree polynomial $f(t)$satisfies $\Delta_\varepsilon^k f(t) = 0$ for $k > n$. Now we rearrange the last sum to obtain the inequality

$$\sum_{j=0}^{N} f\left(\frac{j}{N}\right) \binom{N}{j} \cdot (-)^{N-j} \Delta^{N-j} b_j < 0.$$

But (ii) says $f \geq 0$, and the completely monotone sequence $\{b_i\}$ satisfies $(-)^k \Delta^k b_j \geq 0$, so all terms in the last sum are *nonnegative*, and we have a contradiction. The Farkas alternative (ii) is impossible.

Therefore, the alternative (i) is true: every finite system of moment equations (i) is solvable. It follows that the infinite system (3) is solvable, and so we have proved Hausdorff's theorem.

This theorem is important in probability theory. As William Feller said, "Its discovery has been justly celebrated as a deep and powerful result." (See [**9**], p. 226.)

As you've just seen, the mathematical methods of economics have striking applications to the rest of mathematics. As you might have feared, I could go on talking to you forever. I could tell you about applications to ill-posed boundary-value problems of partial differential equations. But I manfully refrain; you have already heard enough. By now, I hope you will agree with me: these problems and methods of economics are valuable, and they are fascinating.

References

1. Kenneth Arrow, Social Choice and Individual Values, 2nd ed., Yale University Press, New Haven, 1963.

2. Robert G. Bland, The allocation of resources by linear programming, Sci. Amer., 244 (1981) 126–144.

3. Alfred Broaddus, Linear programming: a new approach to bank portfolio management, Monthly Review, Federal Reserve Bank of Richmond, Virginia, vol. 58 (1972) 3–11.

4. Donald J. Brown, Nonstandard economics: a survey, to appear in Studies in Mathematical Economics, S. Reiter, editor, Math. Assoc. of America.

5. George B. Dantzig, Constructive proof of the min-max theorem, Pacific J. Math., 6 (1956) 25–33.

6. _____, Linear Programming and Extensions, Princeton Univ. Press, 1963.

7. Gerard Debreu, Theory of Value, Yale University Press, New Haven, 1959.

8. William Feller, An Introduction to Probability Theory and Its Applications, vol. 1, 3rd ed., Wiley, New York, 1968.

9. _____, An Introduction to Probability Theory and Its Applications, vol. 2, 2nd ed., Wiley, New York, 1971.

10. Joel Franklin, Methods of Mathematical Economics, Springer-Verlag, New York, 1980.

11. John Kenneth Galbraith, Economics, Peace, and Laughter, essays edited by Andrea D. Williams, Houghton Mifflin Company, New York, 1971.

12. O. A. Gross, Preferential arrangements, this MONTHLY, 69 (1962) 4–8.

13. R. F. Harrod, The Life of John Maynard Keynes, Avon Books, New York, 1951.

14. F. Hausdorff, Momentprobleme für ein endliches Intervall, Math. Z., 16 (1923) 220–248.

15. S. Kakutani, A generalization of Brouwer's fixed-point theorem, Duke Math. J., 8 (1941) 457–458.

16. L. V. Khachian, A polynomial algorithm in linear programming. Soviet Math. Dokl., 20 (1979) 191–194 (Amer. Math. Soc. translation).

17. Alan P. Kirman and Dieter Sondermann, Arrow's theorem, many agents, and invisible dictators, J. Econom. Theory, 5 (1972) 267–277.

18. Victor Klee and George Minty, How good is the simplex algorithm? Inequalities, 3 (1972) 159–175, Academic Press, New York.

19. H. W. Kuhn, The Hungarian method for the assignment problem, Naval Res. Logistics Quart., 2 (1955) 83–97.

20. W. A. J. Luxemberg and K. Stoyan, The Theory of Infinitesimals, Academic Press, New York, 1976.

21. G. S. Maddala, Econometrics, McGraw-Hill, New York, 1978.

22. John Milnor, Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem, this MONTHLY, 85 (1978) 521–524.

23. J. F. Nash, Equilibrium points in *n*-person games, Proc. Nat. Academy of Sciences, 36 (1950) 48–49.

24. John von Neumann, Zur Theorie der Gesellschaftsspiele, Math. Ann., 100 (1928) 295–320.

25. John von Neumann and Oskar Morgenstern, Theory of Games and Economic Behavior, Princeton Univ. Press, 1944.

26. H. Nikaido, Convex Structures and Economic Theory, Academic Press, New York, 1968.

27. C. A. Rogers, A less strange version of Milnor's proof of Brouwer's fixed-point theorem, this MONTHLY, 87 (1980) 525–527.

28. Paul Samuelson, Economics, 9th ed., McGraw-Hill, New York 1973.

29. Jacob Schwartz, Mathematics as a tool for economic understanding, pp. 269–295 in Mathematics Today: Twelve Informal Essays, edited by L. A. Steen, Springer-Verlag, New York, 1978.

30. J. A. Sholat and J. D. Tamarkin, The Problem of Moments, Mathematical Surveys, no. 1, American Mathematical Society, New York, 1943.

31. G. F. Shove, The place of Marshall's Principles in the development of economic theory, Econom. J., 52 (1942) 294–329.

32. Eduard Stiefel, An Introduction to Numerical Mathematics, Academic Press, New York, 1963.

33. George Stiefel, The cost of subsistence, J. Farm. Econom., 27 (1945) 303–314.

34. J. J. Sylvester, A question in the geometry of situation, Quart. J. Pure and Appl. Math., 1 (1857) 79.

35. A. Takayama, Mathematical Economics, Holt, Rinehart, and Winston, 1974.
36. Philip Wolfe, The simplex method for quadratic programming, Econometrica, 27 (1959) 382–398.