COWLES COMMISSION DISCUSSION PAPER: ECONOMICS NO. 2108

A Note on Convergence in Linear Programming Problems

Masao Fukuoka

July 14, 1954

In this note we shall consider some systems of differential equations as a means of solving linear programming problems. It should be emphasized that it is rather of a preliminary nature and is intended to summarize various approaches without the rigor of mathematical presentation.

Introduction

Stimulated in large measure by von Mises's criticism, such writers on socialist economy as Lerner and Lange have advocated the use of an automatic market mechanism as a possible computing algorithm for socialist planning. The underlying idea is, as is well known, to set up decentralized pseudo-market processes and, by proceeding with the "game of competition", to let the system gradually approximate the optimum configuration. Indeed, as Professor Koopmans observes [1], if this is relevant to the problems of socialist planning, it will be very natural for economists to apply a similar device to general allocative problems of linear programming. Thus, the following question presents itself: in solving a linear programming problem will there be any workable algorithm which is analogous to the mechanism of market adjustment?

One practical device of computing the solution for linear programming has been known as the "Simplex Method", which may be characterized as a sort of iterative procedure, in each stage making the situation better off and thus stepping up to the optimal goal. Though this is not unlike a convergent market adjustment, explicitly dynamic formulations in terms of differential equations were first provided by Professor Samuelson in his 1949 RAND paper [2]. The present work is originally intended to study a few alternative systems which are free from the difficulties contained in his formulations.

I.

To begin with, let us suppose a typical linear programming problem of maximizing national product: Subject to the constraints

$$(1) \qquad \sum_{j=1}^{n} a_{ij} x_j \leq C_i \qquad (i = 1, 2, \ldots, m)$$

and

$$(2) \qquad x_j \geq 0 \qquad (j = 1, 2, \ldots, n),$$

maximize

$$(3) \qquad Z = \sum_{j=1}^{n} P_j x_j .$$

Here x_j's stand for quantities of various products, P_j's for their prices, a_{ij}'s for fixed technical coefficients, and C_i's for given amounts of various factors. We can also easily write down its dual minimum problem: Subject to

$$(4) \qquad \sum_{i=1}^{m} a_{ij} w_i \geq P_j \qquad (j = 1, 2, \ldots, n)$$

and

$$(5) \qquad w_i \geq 0 \qquad (i = 1, 2, \ldots, m),$$

minimize

$$(6) \qquad Z' = \sum_{i=1}^{m} C_i w_i,$$

where new variables w_i's denote the prices of factors.

Now our problem is this: can we set up any automatic dynamic mechanism which will simultaneously solve both these best price and best quantity problems? In other words, is it possible to design certain dynamic systems whose behavior is in a certain way analogous to that of market adjustment and whose solution will finally converge to a saddlepoint of the following Lagrangean expression

$$V(x,w) = \sum_{j=1}^{n} p_j x_j - \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} w_i x_j + \sum_{i=1}^{m} C_i w_i \qquad (7)$$

As a preliminary to exploring such a problem, Professor Samuelson gave first the following system of differential equations:[1/]

$$(8) \quad \begin{cases} \dfrac{dw_i}{dt} = \alpha_i \left(\sum_{j=1}^{n} a_{ij} x_j - C_i \right) & (i = 1, 2, \ldots, m) \\ \dfrac{dx_j}{dt} = \beta_j \left(p_j - \sum_{i=1}^{m} a_{ij} w_i \right) & (j = 1, 2, \ldots, n) \end{cases}$$

$$w_i \geq 0, \quad x_j \geq 0,$$

where α_i's and β_j's are positive proportionality time constants. This would indeed be the most reasonable procedure for economists to start with, for it describes mathematically the following well-established rules of decentralized market behavior: (i) Whenever the total of demand for any factor exceeds its fixed endowment, let its price, say w_i, increase at a rate proportional to the deficiency. Similarly, when there is a surplus of any factor, let its price decrease at a rate proportional to the surplus -- but with the proviso that, when any w_i goes down to zero, it can fall no further; (ii) Let any output, say x_j, increase at a rate proportional to its profit-

---

1/ See [2], p. 44. System (8) follows immediately by applying the "gradient method" to (7) and putting $\begin{cases} \dot{w} = -\alpha V_w \\ \dot{x} = \beta L_x . \end{cases}$ See [4].

ability. Let x_j decrease if profits are negative at a rate proportional to these losses -- but with the proviso that, when any x_j hits the zero boundary, it cannot go down further. Thus, too low factor prices will yield high profits and hence large expansion of product quantities; this will create excess demand for factors and raise their prices, which will in turn act upon profitabilities adversely and reduce quantities and so forth. The system can only be in equilibrium if profitabilities are everywhere zero for all products whose quantities are not themselves zero, and if surpluses are similarly everywhere zero for all factors whose prices are not themselves zero. For all zero w_i's the corresponding surpluses are nonnegative and for all zero x_j's the corresponding profits are nonpositive.

A crucial defect of this system is that it will not settle down to such an equilibrium. Since the characteristic roots are all imaginary, the w_i's and x_j's will oscillate forever without converging to the optimal solution. Even if some of the variables take zero values, this would still be true for the remaining variables.

II.

This difficulty would possibly be by-passed by several alternative devices, one of which is a device of Professor Samuelson himself. Just like a classical conservative system, the above system oscillates because of its hollow diagonals. Therefore, if we fill them with appropriate non-zero submatrices, we would cause it to converge to the desired optimal configuration. Professor Samuelson's way of doing so is to introduce certain unbiased dissipative or frictional forces. Thus, relying upon this device, he proposed the following corrected equations[2/]

2/ [2], p. 46.

$$
(9) \quad \begin{cases} \dfrac{dw_i}{dt} = \alpha_i \left(\sum_{j=1}^{n} a_{ij}\, x_j - C_i \right) + \sum_{j=1}^{n} K_{ij} \left(P_j - \sum_{i=1}^{n} a_{ij} \cdot v_i \right) x_j & (i=1,2,..,m) \\ \\ \dfrac{dx_j}{dt} = \beta_j \left(P_j - \sum_{i=1}^{n} a_{ij}\, v_i \right) - \sum_{i=1}^{m} M_{ij} \left(\sum_{j=1}^{n} a_{ij}\, x_j - C_i \right) w_i & (j=1,2,..,n) \end{cases}
$$

$$
v_i \geq 0 \quad , n_j \geq 0 \quad ,
$$

where K_{ij}'s and M_{ij}'s are new positive proportionality constants. Economically this might be interpreted to mean that the increase in any factor price should be further intensified by positive profitabilities and similarly the increase in any output be promoted by abundant surplus factors. Through this correction, we are now likely to have the system settle down.

One basic point might be mentioned here as to the mathematical property of the above systems. The proposed sets of differential equations seem to be fairly simple in their linear appearance. However, we should not overlook the crucial effects of nonnegativity boundary inequalities on the variables. As a matter of fact, if even one of the variables is already zero, and by its nature cannot be negative, the corresponding dw_i/dt or dx_j/dt must be equal to either the linear expression or zero whichever is greater. This means that the differential equations are not always appropriate in the linear form and a certain "switching rule" is necessary. Thus, for a satisfactory treatment, we should from the very beginning incorporate the fundamental boundary inequalities into our differential equations by writing them in a certain quasi-linear fashion. (See [4], p. 2).

III

Mr. G. W. Brown and Prof. von Neumann proved in [5] that the optimal mixed strategy for a skew-symmetric zero-sum two-person game is reached as the stationary solution of the following set of quasi-linear differential equations

$$\text{(10)}\qquad \frac{dX_i}{dt} = \varphi\left\{\sum_{j=1}^{k} A_{ij}\, X_j\right\} - X_i \sum_{i=1}^{k} \varphi\left\{\sum_{j=1}^{k} A_{ij}\, X_j\right\} \qquad (i = 1,2,\ldots,k),$$

where X_i's stand for mixed strategies, the array of A_{ij}'s for the skew-symmetric pay-off matrix, and function $\varphi = \varphi\left\{\Sigma A_{ij} X_j\right\} = \varphi\left\{f(X)\right\}$ is defined such that

$$\varphi\left\{f(X)\right\} = \text{Max}\left\{f(X),\; 0\right\}.$$

Since, as is now well known, every linear programming problem can be converted into a symmetric game problem, this technique of Brown and von Neumann may well be applied to the solution of our programming problem.

Considering the specific property of our matrix

$$\text{(11)}\qquad A = \begin{bmatrix} 0 & a & -C \\ -a & 0 & P \\ C & -P & 0 \end{bmatrix}$$

and using the transformation

$$\text{(12)}\qquad \begin{cases} w_i = \dfrac{X_i}{Xm + n + 1} & (i = 1, 2, \ldots, m) \\[2ex] x_j = \dfrac{Xm+j}{Xm + n + 1} & (j = 1, 2, \ldots, n), \end{cases}$$

we can verify that equations (10) can be reduced to the following form[3/]

$$(13)\begin{cases} \dfrac{dw_i}{dt} = \varphi\left\{\sum_{j=1}^{n} a_{ij}\, x_j - C_i\right\} - \varphi\left\{\sum_{i=1}^{m} C_i\, w_i - \sum_{j=1}^{n} P_j\, x_j\right\} w_i & (i=1,2,\ldots m) \\ \dfrac{dx_j}{dt} = \varphi\left\{P_j - \sum_{i=1}^{m} a_{ij}\, w_i\right\} - \varphi\left\{\sum_{i=1}^{m} C_i\, w_i - \sum_{j=1}^{n} P_j\, x_j\right\} x_j & (j=1,2,\ldots,n), \end{cases}$$

where $$\varphi\left\{f(w,x)\right\} = \text{Max}\left\{f(w,x),\ 0\right\}.$$

Note that the boundary conditions are now built into the differential equations themselves. If one of the variables becomes zero, the second term on the right side of the corresponding equation vanishes, and, since function $\varphi\left\{f(w,x)\right\}$ is supposed to be nonnegative, the rate of change of the variable must also be nonnegative. Hence the variable can never fall any further.

One interesting fact to be seen is that the expression $\sum_{i=1}^{m} c_i w_i - \sum_{j=1}^{n} P_j\, x_j$ is the difference between Z', the national cost to be minimized, and Z, the national product to be maximized. As is well known, the basic duality theorem of linear programming guarantees that, in an optimal equilibrium, Z and Z' must be exactly equal to each other, so that all the second terms simultaneously vanish. Thus, we could regard the above expression as an indicator of any discrepancy from the optimum. In equations (13) of the Brown-von Neumann type, this discrepancy will <u>always</u> require downward adjustment to w_i and x_j, while

---

3/ They can also be written as

$$(13)'\begin{cases} \dfrac{dw_i}{dt} = \varphi\left\{\sum_{j=1}^{n} a_{ij}x_j - C_i\right\} - \varphi\left\{-\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j - C_i\right) w_i - \sum_{j=1}^{n}\left(P_j - \sum_{i=1}^{n} a_{ij}w_i\right) x_j\right\} w_i \\ \dfrac{dx_j}{dt} = \varphi\left\{P_j - \sum_{i=1}^{m} a_{ij}w_i\right\} - \varphi\left\{-\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j - C_i\right) w_i - \sum_{j=1}^{n}\left(P_j - \sum_{i=1}^{n} a_{ij}w_i\right) x_j\right\} x_j \end{cases}$$

a discrepancy in each process or market (shown as the first term on the right side of each equation) will <u>always</u> call for upward adjustments of the variables.

Such asymmetry might be avoided if we change slightly the switching rule of our equations so that $\varphi\left\{f(w,x)\right\}$ is now read as

$$\varphi\left\{f(w,x)\right\} = \begin{cases} f(w,x) & \text{for } w,\ x > 0 \\ \mathrm{Max}\left\{f(w,x),\ \ 0\right\} & \text{for } w,\ x = 0\ . \end{cases}$$

Though all the speeds of adjustment are assumed as unity, system (13) with this new switching rule is rather similar to Professor Samuelson's second system (9), and is amenable to an easier market interpretation than system (13) with the original switching rule. Any w_i is now allowed to fall if there is a surplus of that factor, and similarly any x_j can fall if there is a loss for that product -- but, of course, until the corresponding variable hits the zero boundary. The discrepancy between Z and Z' can also act in either direction. However, it should be admitted that, losing the nonnegativity assumption of the φ function, we are now set apart from the proof by Brown and von Neumann and brought back to mere conjecture for the convergence of the system.[4/]

4/ The Brown-von Neumann system was criticized by Mr. Kose for its slow and vibratile convergence. However, as Mr. Bellman suggested, this defect would be easily removed by replacing (10) with

$$(10)' \quad \frac{dX_i}{dt} = \varphi\left\{\sum_{j=1}^{k} A_{ij}\ X_j\right\}^{\alpha} - X_i \sum_{i=1}^{k} \varphi\left\{\sum_{j=1}^{k} A_{ij}\ X_j\right\}^{\alpha},$$

$$\alpha > 0.$$

The smaller the value of α is chosen, the more rapid will be the convergence, and, in the limiting case $\alpha = 0$, it will be even exponential. See [6].

IV.

From the viewpoint of decentralized administration, both the above Brown-von Neumann system and the second Samuelson system seem to have one common disadvantage; they bring forward a tremendous complication into the behavior prescribed for each agency. Every "custodian" must be acquainted not only with the situation of his own market but also with that of all other markets and with profitabilities of all products; similarly every "manager" must know not only the profitability of his own product but also that of all other products and the excess demand for all factors. This would almost be a denial of the idea of decentralization itself.

In view of this disadvantage, it may be desirable to explore the possibility of developing some other systems which are not involved with such complexity. One such system was suggested by Mr. T. Kose of the Otaru Commercial University in Japan. The starting point of his formulation is the following Kuhn-Tucker theorem on concave programming:

> Let the functions $f_1(x), f_2(x), \ldots, f_m(x)$ and $g(x)$ be convave as well as differentiable for all $x \geq 0$. Then the necessary and sufficient condition for x to maximize $g(x)$ subject to $f_1(x) \geq 0, f_2(x) \geq 0, \ldots, f_m(x) \geq 0$, and $x \geq 0$ is that some x^o and w^o satisfy the following conditions for the equation $V(x,w) = g(x) + \sum_{j=1}^{m} w_j f_j(x)$:

$$\begin{cases} \text{(i)} \quad \dfrac{\partial V(x^o,w^o)}{\partial x_j} \leq 0, \quad \dfrac{\partial V(x^o,w^o)}{\partial x_j} x_j^o = 0, \quad x^o \geq 0 \\ \text{(ii)} \quad \dfrac{\partial V(x^o,w^o)}{\partial w_i} \geq 0, \quad \dfrac{\partial V(x^o,w^o)}{\partial w_i} w_i^o = 0, \quad w^o \geq 0 \end{cases}$$

Based upon this theorem, he set forth the system of differential equations of the form

$$(14)\quad \begin{cases} \dfrac{dw_i}{dt} = -\alpha_i \dfrac{\partial V(\varphi(x), w)}{\partial w_i} - \gamma_i \dfrac{d}{dt}\left(\dfrac{\partial V(\varphi(x), w)}{\partial w_i}\right) & (i = 1, 2, \ldots, m) \\[2ex] \dfrac{dx_j}{dt} = \beta_j \dfrac{\partial V(x, \varphi(w))}{\partial w_j} + \delta_j \dfrac{d}{dt}\left(\dfrac{\partial V(x, \varphi(w))}{\partial x_j}\right) & (j = 1, 2, \ldots, n), \end{cases}$$

where α_i, γ_i, β_j, δ_j are positive proportionality constants and

$$\varphi(w) = \mathrm{Max}(w, 0), \quad \varphi(x) = \mathrm{Max}(x, 0).$$

Then he proved that, except when $\lim_{t \to \infty} w$ and $\lim_{t \to \infty} x$ take positive infinite values, $\varphi(\lim_{t \to \infty} w) = w^0$ and $\varphi(\lim_{t \to \infty} x) = x^0$ can be regarded as a solution for the concave programming problem (cf. Reference [7]).

Now, reducing the above system to the form relevant to our linear programming case, we easily get

$$(15)\quad \begin{cases} \dfrac{dw_i}{dt} = \alpha_i \left(\sum_{j=1}^{n} a_{ij}\, \varphi(x_j) - C_i\right) + \gamma_i \dfrac{d}{dt}\left(\sum_{j=1}^{n} a_{ij}\, \varphi(x_j) - C_i\right) & (i=1,2,\ldots,m) \\[2ex] \dfrac{dx_j}{dt} = \beta_j \left(P_j - \sum_{i=1}^{m} a_{ij}\, \varphi(w_i)\right) + \delta_j \dfrac{d}{dt}\left(P_j - \sum_{i=1}^{m} a_{ij}\, \varphi(w_i)\right) & (j=1,2,\ldots,n) \end{cases}$$

where $\varphi(w) = \mathrm{Max}(w, 0), \quad \varphi(x) = \mathrm{Max}(x, 0).$

In this system each agency need no longer pay any attention to others' surpluses and profitabilities, but only to that of his own and its time derivative.

What is interesting to note here is the introduction of the time derivative terms on the right, which is analogous to a device in the servo mechanism. Indeed, if we like, we could regard the rules for decentralized agencies as a kind of reaction characteristics of a communication system (cf. [8]). It is well known that the stability of such system can be improved by introducing derivative or anticipation control as well as displacement or proportional control. In other words, the "error-rate damping" device is as effective means of causing convergence as the "viscous damping device" mentioned in Section II.

References

[1] T. C. Koopmans, "Analysis of Production as an Efficient Combination of Activities", in Activity Analysis of Production and Allocation, 1951.

[2] P. A. Samuelson, "Market Mechanism and Maximization" The RAND Corporation, 1949.

[3] P. A. Samuelson, "Notes on the Dynamic Approach to Saddlepoints and Extremum Points: Part I, Gradient Methods and the Equations of Classical Mechanics", The RAND Corporation, 1951.

[4] L. Hurwicz and K. Arrow, "A Gradient Method for the Lagrangean Problem", Cowles Commission Discussion Paper, 1950.

[5] G. W. Brown and J. von Neumann, "Solutions of Games by Differential Equations", in Contributions to the Theory of Games, Vol. I, 1950.

[6] R. Bellman, "On a New Iterative Algorithm for Finding the Solutions of Games and Linear Programming Problems", The RAND Corporation, 1953.

[7] T. Kose, "Keikaku-ho to Bunkenteki-kettei", Shogaku-tokyu, July, 1953 (with English summary: "Solutions of Programming Problems by Differential Equations").

[8] C. N. Klahr, "Economic Controls Based on Technological Data and Information Theory", Cowles Commission Discussion Paper, 1950.