COWLES COMMISSION DISCUSSION PAPER: ECONOMICS 2053



# A Note on the Optimal Assignment Problem* **

M. Beckmann and T. C. Koopmans

October 30, 1952

1. The problem of optimal assignment of personnel to jobs [Thorndike] or (in a slightly generalized version) of plants to locations may be stated as follows. Given are n persons and n jobs (m plants and n locations, $m \leq n$) and a set of numbers $g_{ij}$ representing the score of person i in job j (profit of plant i at location j). The number $g_{ij}$ is assumed to be independent[1] of the way in which individuals (plants) other than i are assigned to jobs (locations) other than j. What is the assignment of people to jobs (plants to location) that maximizes the total score (profit)?

While nothing in this formulation requires the $g_{ij}$ to be non-negative, any solution for a given $g_{ij}$ is also, as remarked by von Neumann, a solution for $g^*_{ij}$ defined by

$$g^*_{ij} = g_{ij} + u_i + v_j .$$

Without loss of generality we can therefore assume that the $g_{ij}$ are non-negative.

---

* Research undertaken under contract between the Cowles Commission for Research in Economics and the RAND Corporation.

** This paper supersedes Economics 2044 and 2044A which contained errors.

1. In terms of locational analysis such independence of plant income from the locations of other plants means that transportation costs of shipment of intermediate commodities among plants are neglected. Differences of profit of a given plant as between different locations arise from the proximity to localized material deposits, to market sites, and to locations of economical labor supply. [Weber]

Let $P$ be an $n \times n$ permutation matrix,[2/] $G$ the $n \times n$ matrix of the $g_{ij}$ with zeros, in all rows $m+1, \ldots n$, if any. We have

Problem 1: Find $\underset{P}{\text{Max}}$ trace $(GP)$

(or, written individually, $\underset{p_{ij}}{\text{Max}} \sum_{i,j} g_{ij} \; p_{ji}$),

where $g_{ij} \geq 0$.

For easy reference this problem will be denoted as the linear assignment problem.

2. The linear assignment problem can be treated as an activity analysis problem, as is seen by:

Lemma 1: The linear assignment problem is equivalent to the following

Problem 2: Find $\underset{x_{ij}}{\text{Max}} \sum_{ij} g_{ij} x_{ij}$, where $g_{ij} \geq 0$,

subject to the constraints

$$x_{ij} \geq 0$$

$$\sum_j x_{ij} \leq 1$$

$$\sum_i x_{ij} \leq 1$$

Proofs have been given in [von Neumann], and by Dantzig in an unpublished memorandum to Koopmans. The equivalence is stated also in [Votaw and Orden].

---

2. A matrix $P = [p_{ij}]$ obtained from the unit matrix I by a permutation of its columns. More precisely: let $\pi$ be the permutation of integers $i$ to integers $\pi(i)$. Then $P$ is the matrix $[p_{ij}] = [\delta_{i\,\pi(j)}]$, where $\delta_{ik}$ denotes the Kronecker delta. Premultiplication by $P'$ of a matrix permutes the rows of this matrix by the permutation $\pi$:

$$\sum_k p_{ki} \; a_{kj} = \sum_k \delta_{k\,\pi(i)} \; a_{kj} = \cancel{a_{i\,\pi(j)}} \; a_{\pi(i)\,j}$$

Similarly, postmultiplication by $P$ permutes the columns of a matrix by the permutation $\pi$.

The underlying mathematical fact is that the permutation matrices are extreme points of the convex polygon in the space of the $x_{ij}$, defined by the inequalities in problem 2, whereas any other extreme points (obtained from the permutation matrices by replacing some ones by zeros) can give no higher values of the maximand whenever $g_{ij} \geq 0$.

Now problem 2 is a linear activity analysis problem (treating persons or plants and jobs or localities as perfectly divisible instead of indivisible). From the efficiency price theorem [5.4.1, Koopmans] (applied with $y = \Sigma g_{ij} x_{ij}$ as the only final commodity outflow, the variables $z_i = -\sum_j x_{ij}$ and $w_j = -\sum_i x_{ij}$ as inflows of primary commodities, and the variables $x_{ij}$ as flows of intermediate commodities) we conclude the existence of non-negative parameters $\lambda_i$, $\mu_j$, $\nu_{ij}$ such that

$$(0) \qquad g_{ij} = \lambda_i + \mu_j - \nu_{ij} \qquad \text{with } \nu_{ij} = 0 \quad \text{if } \bar{x}_{ij} > 0,$$

where $\bar{x}_{ij}$ represents a solution $x_{ij}$ of Problem 2. On the other hand it follows from Lemma 1 that $\bar{x}_{ij}$ is (or at any rate can be chosen as) a permutation matrix (that is, in agreement with the actual indivisibility of person or plants and jobs or locations) $\bar{x}_{ij} = \delta_{\pi(i)j}$ where $\delta_{ij}$ is the Kronecker delta and $\pi(i)$ denotes the index obtained by applying the permutation $\pi$ to i (Cf. footnote p. 2). Hence if j is any index denoting a location

$$(1) \qquad g_{ij} \begin{Bmatrix} = \\ \leq \end{Bmatrix} \lambda_i + \mu_j \quad \text{for } j \begin{Bmatrix} = \\ \neq \end{Bmatrix} \pi(i)$$

Incidentally, it follows from the duality theorem [Dantzig; Gale, Kuhn and Tucker] that $\lambda_i$, $\mu_j$ is the solution of

$$\underset{\lambda_i, \mu_j}{\text{Min}} \left( \sum_i \lambda_i + \sum_j \mu_j \right)$$

subject to

$$g_{ij} \leqq \lambda_i + \mu_j$$
$$0 \leqq \lambda_i$$
$$0 \leqq \mu_j.$$

These statements may be summarized in the following

<u>Lemma 2</u>

Let $P = [\delta_{\pi(i)j}]$ be a solution of the linear assignment problem. Then $\bar{x}_{ij} = \delta_{\pi(i)j}$ is the solution of

$$\underset{x_{ij} \geqq 0}{\text{Min}} \quad \underset{\lambda_i \geqq 0, \mu_j \geqq 0}{\text{Max}} \qquad \phi(..x_{ij}..,..\lambda_i..,..\mu_j..)$$

where
$$\phi = \sum_{ij} g_{ij}x_{ij} + \sum_i \lambda_i (1 - \sum_j x_{ij}) + \sum_j \mu_j (1 - \sum_i x_{ij})$$

and
$$\underset{x_{ij} \geqq 0}{\text{Min}} \quad \underset{\lambda_i \geqq 0, \mu_j \geqq 0}{\text{Max}} \ \phi \quad = \underset{\lambda_i \geqq 0, \mu_j \geqq 0}{\text{Max}} \quad \underset{x_{ij} \geqq 0}{\text{Min}} \ \phi$$

The sufficiency of condition (1) must be understoodas being subject to the constraint that each plant has to be combined with some location.

3. Equation (1) has an interesting economic interpretation. Assume that rents $\ell_j$ and $p_i$ are levied on locations $j$ and plants $i$, respectively. A set of these rents $\ell_j$, $p_i$ shall be called a set of <u>efficiency rents</u> if it has these properties: 1) For each plant $i$, income $g_{ij} - \ell_j$ after rent on location is maximized (among all locations $j$) at the location $j = \pi(i)$ which is assigned to the plant by the solution of the optimum assignment problem.

(2) For each location j, income $g_{ij} - p_i$ after rent on plant is maximized (among all plants i) by that plant (i such that $\pi(i) = j$) which is assigned to this location by the solution of the optimum assignment problem. 3) For each combination of a plant i and a location j, income $g_{ij} - p_i - \ell_j$ after rent on both plant and location is at most zero, and equals zero if the combination (i,j) occurs in the optimal assignment $j = \pi(i)$.

Equation (1) asserts that there exists a set of efficiency rents, that

$$p_i = \lambda_i$$

$$\ell_j = \mu_j$$

constitutes such a set of efficiency rents, and that $p_i + \ell_j$ exhausts the maximal income $g_{ij}$. This imputation of profits to locations and plants is, however, not necessarily unique. Even if there is only one solution to problem 2, the rents $\mu_j$ are not generally determined in that they are free to vary over a range set by the numerically nearest assignments. That is to say if $\lambda_i$, $\mu_j$ is a set of efficiency rents, so is

$$(2) \qquad \begin{cases} \lambda_i^* = \lambda_i + \epsilon + \epsilon_i \\ \mu_j^* = \mu_j - \epsilon - \epsilon_{\rho(j)} \end{cases}$$

where $\mathcal{E}$ is unrestricted as to sign or magnitude, where $\rho$ is the permutation $\pi^{-1}$ inverse to $\pi$, ($i = \rho(i)$ whenever $j = \pi(i)$) and where the $\epsilon_i$ are restricted by the inequalities in (1),

$$(3) \qquad g_{ij} \leq \lambda_i^* + \mu_j^* = \lambda_i + \mu_j + \epsilon_i - \epsilon_{\rho(j)} \text{ for } i \neq \rho(j)$$

or, from (0),

$$(4) \qquad \mathcal{E}_{\rho(j)} - \mathcal{E}_i \leq \nu_{ij} \qquad \text{for } i \neq \rho(j).$$

The number $\epsilon$ represents a uniform transfer of rent income from locations to plants. Its indeterminacy indicates that our problem does not imply anthing about absolute rent levels on plants and locations, but only restricts to some extent rent differences between the several plants on the one hand, and between locations on the other. The meaning of these restrictions is shown by (4)

to be that any redistribution of rents among plants, accompanied by a compensating redistribution of rents among their partner locations in the optimum assignment, is permissible as long as no non-optimal plant-location combination would thereby be made to render a positive net revenue.

On the other hand, the equations (2) exhaust all transformations of the rents $\lambda_i$, $\mu_j$ associated with a given optimal assignment $\pi$, which conserve the optimality conditions (1). The limited indeterminacies found for the $\epsilon_i$ are connected with the fact that plants and locations occur in our problem as indivisible entities.

Efficiency rents can be used to achieve a decentralization of the assignment decisions, in the following way.

Since the net income of each plant is maximized at that location that is assigned to the plant by a solution of the linear assignment problem, no entrepreneur can benefit from a locational change out of the optimum assignment. Conversely, since the condition (1) is sufficient for a solution of the assignment problem, every combination of plants and locations in which each plant find its income maximized represents a solution of the optimum assignment problem. The reward of a maximal income at the (socially) optimal location, or rather the penalty of a negative net income at all non-optimal locations, is therefore a sufficient means for maintaining in existence an assignment of plants to location which represents a solution of the linear assignment problem.

4. von Neumann has shown that the personnel assignment problem is equivalent to the following game. Let there be n x n double indexed cells, say fields

in a square matrix. Player I hides in one cell. Player II attempts to "find" I by guessing either of the indices of the cell in which player I has hidden, announcing which kind of index (row or column) he is guessing. The payoff to player II is $1/g_{ij}$ if the cell $(i,j)$ in which I is found is in the row or column specified and zero otherwise. The optimal strategies are then for player I to hide in cell ij with probability $\delta_{\pi(i)j}$ and for player II to guess row $i$ or column $j$ with probabilities proportional to $\lambda_i$ and $\mu_j$. Here the $\lambda_i$, $\mu_j$ are subject to the constraints imposed by equation (1) and by the non-negativity condition $\lambda_i \geqq 0$ $\mu_j \geqq 0$. (Non-negativity was seen to be irrelevant for efficiency rents, but is important here.) To the extent that $\frac{\lambda_i}{\sum_i \lambda_i + \sum_j \mu_j}$ and $\frac{\mu_j}{\sum_i \lambda_i + \sum_j \mu_j}$ are indeterminate, so is the optimal strategy of player II.

5. We shall now state a problem that arises as a natural generalization of the linear locational assignment problem. Let there be r plants and n locations, $n \geqq r$; and let a set of non-negative numbers $a_{k\ell}$ represent the commodity flows (in weight units) from the $k^{th}$ plant to the $\ell^{th}$ plant, and a set of positive numbers $b_{ij}$, $i \neq j$, the cost of transportation for the unit flow from location $i$ to location $j$. Assume that the $a_{k\ell}$ are independent of the locations assigned, and that the transport costs $b_{ij}$ are independent of the amounts and compositions of all flows.

Let there be a third set of non-negative numbers $c_{ki}$ denoting the gross revenue of plant k at location i before rent and transportation costs for intermediate commodities (those obtained from or supplied to other plants). What is the assignment of plants to locations which maximizes the sum of all gross revenues less total cost of transportation among plants? This problem will be referred to as the quadratic assignment problem.

As a special case suppose that $n = r$, $c_{ik} \equiv 0$ and that flows from the $k^{th}$ plant all go to the $k + 1^{th}$ plant, those from the $n^{th}$ plant flowing back to the first plant, and that these flows are all equal. In this case we have the familiar travelling salesman problem [Robinson].

The quadratic assignment problem may also be stated, with a grain of salt, in terms of personnel placement in teams. Let there be n persons and n jobs, and let $b_{ij}$ be an index of "compatibility" of persons i and j and $a_{k\ell}$ the amount of collaboration, suitably measured, that is required between the $k^{th}$ and $\ell^{th}$ job, and let $c_{ki}$ be the score of person k in the $i^{th}$ job. What is the assignment of people to jobs that will maximize the performance of the team?

A mathematical formulation is obtained as follows. Let $r + 1, \ldots\ n$ denote "dummy plants" at which no flows originate or terminate, $a_{r+1\ell} = \cdots a_{n\ell} = 0$ $a_{k\ r+1} = \cdots = a_{kr} = 0$. Let A be the $n \times n$ matrix of $a_{k\ell}$, B be the $n \times n$ matrix of $b_{ij}$, putting $b_{ii} = 0$. Finally, let C be the $n \times n$ matrix of $c_{ki}$. Denote by P a $n \times n$ permutation matrix (cf. footnote 2).

<u>Problem 3</u> Find $\underset{P}{\text{Max}}$ trace $(CP - A'P'BP)$.

The case of the travelling salesman problem is obtained when $C = 0$ and A is itself a permutation matrix of the form $\begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \ddots \\ 0 & 0 & & 1 \\ 1 & 0 & & 0 \end{pmatrix}$.

6. If transportation costs are independent of the direction, $b_{ij} = b_{ji}$, both matrices A and B can be assumed symmetric and a statement of the problem be made which brings out more clearly its quadratic nature. Let e, f be real numbers. Then

$$\begin{aligned} & \operatorname{tr} \left\{ (A + eI)'P'\ (B + fI)P \right\} \\ = & \operatorname{tr} \left\{ (A'P'BP) + eP'BP + fA'P'P + efP'P \right\} \\ = & \operatorname{tr}(A'P'BP) + e \operatorname{tr}(BPP') + f \operatorname{tr}(A'P'P) + ef \operatorname{tr}(P'P) \\ = & \operatorname{tr}(A'P'BP) + e \operatorname{tr} B + f \operatorname{tr} A + ef \operatorname{tr} I \end{aligned}$$

Hence the solution to our problem is not changed if we replace A and B by new matrices.

$$\hat{A} = A + eI \qquad \hat{B} = B + fI$$

By choosing e and f large and positive we can make both $\hat{A}$ and $\hat{B}$ positive definite. Since $\hat{A}$ and $\hat{B}$ are symmetric there exist then symmetric positive definite square roots

$$\hat{A} = EE = EE' \qquad \hat{B} = FF = F'F$$

Now
$$\mathrm{tr}(\hat{A}'P'\hat{B}P) = \mathrm{tr}\,(EE'P'F'FP)$$
$$= \mathrm{tr}\,(E'P'F'FPE)$$
$$= \mathrm{tr}\,\{(FPE)'\,(FPE)\}$$
$$= \sum_{j\ell}\,(\sum_{ik} f_{ij}\,p_{ik}\,e_{k\ell})^2 = \mathrm{norm}\,(FPE)$$

The quadratic assignment problem may then be written as

Problem 4 Find a permutation matrix P such that for a given non-negative r x r matrix C and for given, symmetric, positive n x n matrices E, F

$$\mathrm{trace}\,(CP) - \mathrm{norm}\,(FPE)$$

is a maximum.

7. As in the case of the linear assignment problem let us attempt to define efficiency rents. Since now the costs of transportation for intermediate commodities enter explicitly it will be necessary also to introduce efficiency prices on the locations of these commodities. Let the double index $k\ell$ denote the commodity (or aggregate of commodities) that is shipped from plant k to $\ell$. We shall use notations

| | |
|---|---|
| $p_k$ | for rent on plant k |
| $\ell_i$ | for rent on location i |
| $q_{k\ell,i}$ | for the price of commodity $k\ell$ at location i. |

Let $\sigma$ be an assignment of locations to plants $i = \sigma(k)$, that is, $\sigma$ is a permutation.

Net revenue after rents and after charges and returns on intermediate commodities of plant k at location $i = \sigma(k)$ is now

$$c_{k\sigma(k)} - p_k - \ell_{\sigma(k)} + \sum_{\ell} (q_{k\ell,\sigma(k)}\, a_{k\ell} - q_{\ell k,\sigma(k)}\, a_{\ell k})$$

We now consider the question whether the $\ell_i$ and $p_k$ can be regarded as efficiency rents and the $q_{k\ell,i}$ as efficiency prices. Our definition of efficiency prices (p. 4) would require that, for any assignment $i = \sigma(k)$,

(5) $$c_{k\sigma(k)} - p_k - \ell_{\sigma(k)} + \sum_{\ell} (q_{k\ell,\sigma(k)}\, a_{k\ell} - q_{\ell k,\sigma(k)}\, a_{\ell k}) \leqq 0,$$

$k = 1, \ldots n$, and that the "=" sign holds for all k if $\sigma = \pi$, where $\pi$ is the permutation which represents a solution of the assignment problem. The notion of an efficiency price on a commodity location implies that

(6) $$q_{k\ell,j} - q_{k\ell,i} \leqq b_{ij} \qquad k, \ell, i, j = 1, \ldots n$$

and that the "=" sign holds if there is a positive flow of the commodity k from i to j.

Presently we show that with these properties the efficiency rents and efficiency prices $p_k$, $\ell_i$, $q_{k\ell,i}$ satisfy the requirements for efficiency prices on the primary and intermediate commodities of a certain linear activity analysis problem.

Denote by $x_{ki}$ the level of the "activity": combining plant k with the location i, and by $x_{k\ell,ij}$ the "activity level" of flows of commodity $k\ell$ from i to j. For the moment we disregard the fact that, because of the indivisibility of plants, each $x_{ki}$ is either 0 or 1, and each $x_{k\ell,ij}$ either 0 or $a_{k\ell}$. We define a final commodity flow (net revenue)

$$y_0 \equiv \sum_{ki} c_{ki}\, x_{ki} - \sum_{k\ell,ij} b_{k\ell,ij}\, x_{k\ell,ij},$$

and intermediate commodity flows

$$y_{k\ell,i} = \sum_j (x_{k\ell,ij} - x_{k\ell,ji}) - a_{k\ell} x_{ki} + a_{k\ell} x_{\ell i}$$

and impose the following availability constraints on primary and intermediate commodities

$$y_i = \sum_k x_{ki} \leq 1, \qquad y_k \sum_i x_{ki} \leq 1, \qquad y_{k\ell,i} \geq 0$$

Then it can be shown in straight forward manner that the $p_k$, $\ell_i$, $q_{k\ell,i}$ are the efficiency prices associated with the primary commodities of which the flow is measured by $y_k$, $y_i$ and the intermediate commodities measured by $y_{k\ell,i}$ respectively.

Now it is easy to see from examples that the solution of this activity analysis problem represents in general not an extreme point of the set of possible $x_{ki}$; that is to say, the solutions $\bar{x}_{ki}$ of this problem are not in general of the form $\bar{x}_{ki} = \delta_{i\sigma(k)}$, where $\sigma$ is a permutation. For instance if $c_{ki} \equiv 1$ then the solution is easily found to be

$$\bar{x}_{ki} = \frac{1}{r} \qquad \bar{x}_{k\ell,ij} = \begin{Bmatrix} 0 \\ \frac{1}{r} \end{Bmatrix} \text{ if } i \begin{Bmatrix} \neq \\ = \end{Bmatrix} j \qquad k, \ell, i, j = 1, \ldots n$$

and therefore is not an extreme point. Changes in the parameters $c_{ki}$, unless $c_{ki} = \delta_{\tau(k)i}$ with $\tau$ a permutation, will not reverse this occurrence of mixtures of plants in each location, once the indivisibility requirement ($x_{ki} = 0$ or 1) is dropped. Now the conditions (5)&(6) are sufficienct for a solution of the linear activity analysis problem just described. Hence if there existed efficiency rents and efficiency prices associated with a solution of the quadratic assignment problem, satisfying the requirements (5), (6), this would imply that

$$\bar{x}_{ki} = \delta_{\rho(k)i}$$

is a solution of the associated linear activity analysis problem. But this was seen to be not true in the general case. We conclude that in general no efficiency prices satisfying (5) and (6) can be associated with the quadratic locational assignment problem.

## References

Dantzig, G. B., "A Proof of the Equivalence of the Programming Problem and and the Game Problem.", Activity Analysis of Production and Allocation , (T. C. Koopmans, Ed.), New York, 1951. pp. 330-335, especially p. 332.

Gale, D., Kuhn, H. W., and Tucker, A. W. "Linear Programming and the Theory of Games", Activity Analysis of Production and Allocation, pp. 317-329, especially p. 322, theorem 2.

Koopmans, T. C., "Analysis of Production as an Efficient Combination of Activities", Activity Analysis of Production and Allocation, (T. C. Koopmans, Ed.), New York, pp. 33-97.

Robinson, Julia, "On the Hamiltonian Game ( A Travelling Salesman Problem)", RAND paper, December 5, 1949.

Slater, M."Lagrangean Multipliers Revisited", Cowles Commission Discussion Paper, Mathematics 403.

Thorndike, R. L., "The Problem of Classification of Personnel", Psychometrika, Volume 15 (1950), pp. 215-235.

von Neumann, J.,"The Problem of Optimal Assignment and a Certain 2-Person Game." unpublished manuscript, October 26, 1951.

Votaw, D. F. and Orden, A., "The Personnel Assignment Problem", Project Scoop, Symposium on Linear Inequalities and Programming, No. 10, April, 1952, pp. 155 - 163.

Weber, Alfred,"Theory of Location of Industries" ( C. J. Friedricks, ed.), Chicago, 1929, especially pp. 41, 95.