A THEOREM ON CONDITIONAL EXTREMES WITH AN APPLICATION TO TOTAL DIFFERENTIALS

L. M. COURT

1. Introduction. In a paper [1] published in 1941, the writer showed that if one function is maximized subject to a linear constraint, a related function is minimized subject to an associated linear constraint. Specifically he showed that if \( \phi(q_1, \ldots, q_n) \) is maximized subject to the “single” constraint \( \sum_{i=1}^{m} p_i q_i = m \) and \( q_i = q_i(p_1, \ldots, p_n) \) \((i=1, \ldots, n)\) are the first order conditions for this maximization, then the function \( \psi(p_1, \ldots, p_n) = \phi[q_1(p_1, \ldots, p_n), \ldots, q_n(p_1, \ldots, p_n)] \) is minimized subject to the same constraint except that the \( p \)'s are now regarded as the active variables. In the paper mentioned, he gave geometric and analytic proofs that the first order conditions for \( \psi \) are satisfied, but attempted only a geometric proof for the second order conditions.

In the present paper, this theorem is extended to the case in which the constraint is no longer bilinear in the two sets of variables but quite general. Moreover, we admit more than one constraint, in fact any number fewer than the number of \( q \) (or \( p \)) variables. The proof is analytic for the second order conditions as well as the first. Finally, in a supplementary section, it is shown how the two-way relationship between \( \phi, \psi \) and the constraints sometimes can assist in the integration of total differentials.

The present theorem bears more than a passing resemblance to Mayer’s law of reciprocity for isoperimetric problems in the calculus of variations [2]. There the same extremals that maximize or minimize one integral subject to a second, also maximize or minimize the second subject to the first.

2. The theorem. Suppose that \( \phi(q_1, \ldots, q_n) \) and \( \phi(q_1, \ldots, q_n; p_1, \ldots, p_n) (\alpha = 1, \ldots, m; m < n) \) and their first and second partial derivatives are defined and continuous for \( Q = (q_1, \ldots, q_n) \in \mathcal{A} \) and \( P = (p_1, \ldots, p_n) \in \mathcal{B} \), where \( \mathcal{A} \) and \( \mathcal{B} \) are domains in their respective spaces. Suppose furthermore that there is a one-to-one correspondence between the points of \( \mathcal{A} \) and \( \mathcal{B} \); that is, to each \( Q \in \mathcal{A} \), there is a \( P = (p_1, \ldots, p_n) \equiv (p_0^0, \ldots, p_n^0) \in \mathcal{B} \) such that \( \phi(q_1, \ldots, q_n) \) is rendered stationary at \( Q \) subject to the \( m \) constraints \( \phi(q_1, \ldots, q_n; p_1, \ldots, p_n) = 0 \), the \( p \)'s being regarded as parameters or inactive.
variables in these constraints. Let \( q_{t} = q_{i}(p_{1}, \ldots, p_{n}) \) \((i = 1, \ldots, n)\) be the relations that render \( \phi \) conditionally stationary at \( Q \)—they are the relations that express the correspondence between the points of \( \mathcal{A} \) and \( \mathcal{B} \). Then these relations \( q_{t} = q_{i}(p_{1}, \ldots, p_{n}) \)—or, what amounts to the same thing, their inverses \( p_{K} = p_{K}(q_{1}, \ldots, q_{n}) \) \((K = 1, \ldots, n)\)—will also render the related function \( \psi(p_{1}, \ldots, p_{n}) \) stationary at \( P \) subject to the \( m \) constraints \( \alpha G(q_{1}, \ldots, q_{n}; p_{1}, \ldots, p_{n}) = 0 \), the \( q \)'s now being regarded as the inactive variables. Moreover, if the second order conditions in Lagrange multiplier form \([3]\) for the proper or improper maximization (minimization) of \( \phi(q_{1}, \ldots, q_{n}) \) subject to \( \alpha G(q_{1}, \ldots, q_{n}; p_{1}, \ldots, p_{n}) = 0 \) \((\alpha = 1, \ldots, m)\), the \( p \)'s inactive, are satisfied at \( Q \), then the second order conditions in the same form for the proper or improper minimization (maximization) of \( \psi(p_{1}, \ldots, p_{n}) \) subject to the constraints \( \alpha G(q_{1}, \ldots, q_{n}; p_{1}, \ldots, p_{n}) = 0 \), the \( q \)'s inactive, will be satisfied at \( P \).

3. Notation. The general \( q \) variable will always carry a lower case (small) subscript, the general \( p \) variable a capital subscript—\( q_{i}, q_{t}, p_{K}, p_{s} \); this convention will be all the more adhered to in summations. Differentiation with respect to a \( q \) variable will be denoted by a lower case subscript, with respect to a \( p \) variable by a capital subscript: \( \phi_{i} = \partial \phi / \partial q_{i}, \psi_{K} = \partial \psi / \partial p_{K}, \; \alpha G_{i} = \partial \alpha G / \partial q_{i}, \; \alpha G_{K} = \partial \alpha G / \partial p_{K}, \; \alpha \lambda_{i} = \partial \alpha \lambda / \partial q_{i}, \; \psi_{iK} = \partial \psi_{K} / \partial p_{K}, \; \phi_{ij} = \partial^{2} \phi / \partial q_{i} \partial q_{j}, \; \psi_{KL} = \partial^{2} \psi / \partial p_{K} \partial p_{L}, \; \alpha G_{iK} = \partial \alpha G / \partial q_{i} \partial p_{K}, \) and so forth. The umbral notation will be used, and a repeated subscript will mean summation over the natural range of the subscript—1 to \( m \) for \( \alpha \) and 1 to \( n \) for all other subscripts.

4. The first order conditions for \( \psi \). The fact that \( \phi(q_{1}, \ldots, q_{n}) \) is stationary at \( Q \) subject to the constraints

\[
\alpha G(q_{1}, \ldots, q_{n}; p_{1}, \ldots, p_{n}) = 0 \quad (\alpha = 1, \ldots, m)
\]

where \( P = (p_{1}, \ldots, p_{n}) \subseteq \mathcal{B} \) (this is assumed) can be expressed in the Lagrange multiplier form

\[
\phi_{i} - \alpha \lambda \alpha G_{i} = 0 \quad (i = 1, \ldots, n);
\]

these equations are, of course, necessarily satisfied when \( \phi \) is a conditional extreme at \( Q \). The second order conditions for the conditional extremization of \( \phi \), in Lagrange multiplier form, when \( \phi \) is conditionally extremized (\( \phi \) can be stationary at a point without being extremized), hinge critically upon the conditional quadratic form in the \( \xi \)'s.
A THEOREM ON CONDITIONAL EXTREMES 425

(3) \[ \{ \phi_{ij} - a\lambda \, aG_{ij} \} \xi_i \xi_j \]

this quadratic form being subject to the linear constraints (again in the \(\xi\)’s)

(4) \[ aG_{i} \xi_i = 0 \quad (\alpha = 1, \ldots, m). \]

\(\phi\) will be properly or improperly maximized (minimized) subject to its constraints according as this quadratic form is conditionally negative (positive) definite or semidefinite.

(1) and (2) can be regarded as equations in the “unknowns” \(a\lambda\) and \(q_i\); we know that if \(a\lambda^p\) and \(q_i^p\) are a particular set of values for which these equations are satisfied, then they can be solved to give the \(a\lambda\) and \(q_i\) as unique, single-valued functions of the \(p_K\) within some neighborhood of \(P = (p_1^0, \ldots, p_n^0)\), the values of the functions (regarded as a vector) lying within some neighborhood of the augmented point \((\lambda^p, \ldots, m\lambda^p; q_1^p, \ldots, q_n^p)\), provided the jacobian of the left-hand members of (1) and (2) with respect to \(a\lambda, q_i\) does not vanish at \((\lambda^p, \ldots, m\lambda^p; q_1^p, \ldots, q_n^p; p_1^0, \ldots, p_n^0)\). The jacobian in question is nothing other than the bordered determinant

\[
\begin{vmatrix}
0 & \cdots & 0 & \frac{a\lambda}{aG_{i1}} & \cdots & \frac{a\lambda}{aG_{in}} \\
0 & \cdots & 0 & \frac{m\lambda}{aG_{i1}} & \cdots & \frac{m\lambda}{aG_{in}} \\
\frac{aG_{i1}}{aG_{11}} & \cdots & \frac{aG_{i1}}{aG_{1n}} & 0 & \cdots & 0 \\
\frac{aG_{in}}{aG_{n1}} & \cdots & \frac{aG_{in}}{aG_{nn}} & 0 & \cdots & 0 \\
\end{vmatrix}
\]

associated with the conditional quadratic form (3), (4), the jacobian and the determinant both being different from zero whenever \(\phi\) is properly conditionally maximized or minimized. At all other points \(Q \in A\), the possibility of expressing the \(q_i\) (and the associated \(a\lambda\)) in terms of the \(p\)'s is guaranteed by an explicit assumption in the hypothesis of the theorem.

By definition,

\[ \psi(p_1, \ldots, p_n) = \phi[q_1(p_1, \ldots, p_n), \ldots, q_n(p_1, \ldots, p_n)] \]

and \(\psi_K = \phi_i q_i K\); also, if we think of the \(q_i\)'s in (1) as functions of the \(p_i\)'s, \(aG_{i} q_i K + aG_K = 0\). If now we multiply (2) through by \(q_i K\) and use the umbral notation to sum over all \(i\),

\[ 0 = \phi_i q_i K - a\lambda \, aG_{i} q_i K = \psi_K - (-a\lambda) \, aG_{i} K \]

\[ \equiv \psi_K - a\sigma \, aG_{i} K \quad (K = 1, \ldots, n), \]
where the $\lambda = -\lambda$ are new Lagrange multipliers. (5), of course, implies that $\psi$ is stationary at $P$ subject to the constraints $\partial G(q', \cdots, q'_n; p_1, \cdots, p_n) = 0$, the $q'$s being regarded as inactive.

5. The conditional extremization of $\psi$. We can think of the $p'$s in (2) as replaced by their values in terms of the $q'$s—by hypothesis $B$ is mapped in a one-to-one manner onto $A$ so that not only is every $q_i$ a single-valued function of the $p'$s but conversely every $p_K$ is a single-valued function of the $q'$s—and then differentiate this equation with respect to $q_j$:

$$\phi_{ij} - \lambda \partial G_{ij} = \lambda aG_{ij} + \lambda aG_{ij} p_{ij} = 0 \text{ or }$$

$$\phi_{ij} - \lambda \partial G_{ij} = \lambda aG_{ij} + \lambda aG_{ij} p_{ij} \quad (i, j = 1, \cdots, n).$$

If the last equation is multiplied through by $\xi_i \xi_j$, summed with respect to $i$ and $j$, and (4) is taken into account, we recognize the equivalence of the two quadratic forms in the $\xi$'s:

$$\{\phi_{ij} - \lambda aG_{ij}\} \xi_i \xi_j = \lambda aG_{ij} p_{ij} \xi_i \xi_j.$$

That is, either one of these forms, taken subject to the constraints (4), is positive or negative definite (or indefinite) if the other is.

Clearly there must be a similar equivalence between the two parallel quadratic forms in the $\eta$'s,

$$\{\psi_{KL} - \sigma aG_{KL}\} \eta_K \eta_L = \sigma aG_{KL} q_{KL} ^{\eta_K \eta_L}$$

subject to the parallel linear constraints

$$aG_K \eta_K = 0 \quad (\alpha = 1, \cdots, m),$$

these linear and quadratic forms being related to $\psi$ and the $aG'$s, regarded as functions of the active variables $p_K$, in the same way as the quadratic forms in (7) and the linear constraints (4) are related to $\phi$ and the $aG'$s, regarded as functions of the $q_i$.

(In the hypothesis of our theorem we have assumed the existence of only the first and second partial derivatives of $\phi$ and the $aG$, and some doubt may arise about whether this is sufficient to insure the existence of the $\psi_{KL}$ in (8). Now $\psi(p_1, \cdots, p_n) \equiv \phi[q_1(p_1, \cdots, p_n), \cdots, q_n(p_1, \cdots, p_n)]$ and $\psi_{KL} = \phi(q_K q_L + q_K q_L \cdot \cdot \cdot q_K q_L)$. But the $q'$s are obtained as functions of the $p'$s by solving (1) and (2), so that at first sight the existence of the $q_{KL}$, and therefore $\psi_{KL}$, would seem to depend on the existence of the third partials $\phi_{ijl}$. However, $\psi_{KL} = aG_{KL} + aG_{KL} + aG_{KL} q_{KL}$ (this equation is a strict parallel to (6) and is derived in the same way), and since $\sigma = -\lambda$, $\psi_{KL}$ exists if the second partials of $\phi$ and the $aG$ do.)
From (1), \( \partial G q_i u = - \partial G u \), so that if we set

\[
\xi_i = q_i w \eta v \\
\partial G_i \xi_i = \partial G_i q_i w \eta v = - \partial G u \eta v = 0 \quad (\alpha = 1, \ldots, m),
\]

the final equation being true in virtue of (9); and since the constraints (4) are satisfied—in fact, the real meaning of (11) is that (4) and (9) are equivalent—the values assigned to the \( \xi_i \) by (10) are legitimate. If now we introduce (10) into the second form in (7) and remember that \( \partial q_i w \eta v = \delta_{sv} \) (the Kronecker delta), we get

\[
\lambda \partial G_{is} \partial q_i w \eta v q_i w \eta v = \lambda \partial G_{is} (\partial q_i w \eta v q_i w \eta v)
\]

\[
= \lambda \partial G_{iv} \delta_{sv} \eta v q_i w \eta v = \lambda \partial G_{iv} q_i w \eta v \eta v
\]

\[
= - \lambda \partial G_{iv} q_i w \eta v \eta v,
\]

the final quadratic form being (because of the assumed continuity of the second partial derivatives of the \( \partial G \)) the same except for the minus sign as the one on the right in (8). In other words, under the linear constraints (4) or their equivalent (9), either quadratic form in (8) is the negative of the corresponding quadratic form in (7), and vice-versa.

From this the second part of our theorem, having to do with the reciprocal relationship between the conditional maxima and minima of \( \phi \) and \( \psi \), follows immediately.

6. **Application to total differentials.** The result that \( \psi \) is conditionally stationary whenever \( \phi \) is can be applied to total differentials. Let the coefficients \( C_i \) of the differential expression

\[
C_i(q_1, \ldots, q_n) dq_i = [\gamma(q_1, \ldots, q_n)]^{-1} \phi_i dq_i
\]

satisfy \( (n-1)(n-2)/2 \) independent six-term integrability conditions, \( \phi(q_1, \ldots, q_n) \) being a parent integral. Suppose we can find \( m \) functions \( \partial G(q_1, \ldots, q_n; p_1, \ldots, p_n) \) such that the \( m+n \) equations

\[
\partial G = 0, \quad C_i - \partial \mu \partial G_i = 0 \quad (\alpha = 1, \ldots, m; i = 1, \ldots, n)
\]

can be solved to give the \( m+n \) "unknowns" \( \partial \mu, q_i \) as unique single-valued functions of the \( p_K \), the point \( (q_1, \ldots, q_n) \) ranging over some domain \( A \) as \( (p_1, \ldots, p_n) \) ranges over a domain of its own \( B \). What this really means is that the parent integral \( \phi \) is being rendered stationary subject to \( \partial G = 0 \) for \( (p_1, \ldots, p_n) \in B \). Since \( \gamma(C_i - \partial \mu \partial G_i) \equiv \phi_i - \partial \mu \partial G_i \), the \( \partial \mu \) can be identified with the \( \lambda \) in (2); and we see that the values of the \( q_i \) in terms of the \( p_K \) obtained by solving (13) are the same as those obtained by solving (1) and (2) if the \( \phi \) and...
the $aG$ of the present section are identified with those of the earlier sections. Moreover,

$$C_i q_{iK} = a\mu aG_i q_{iK} = -a\mu aG_K$$

$$= -\gamma^{-1} a\lambda aG_K = \gamma^{-1} a\sigma aG_K = \gamma^{-1} \psi_K,$$

so that the total differential

$$DK(p_1, \ldots, p_n)dp_K = -a\mu aG_K dp_K = \gamma^{-1} \psi_K dp_K$$

is integrable. (The $D_K$ in (15) are defined by the identity $D_K \equiv -a\mu aG_K$, in which we think of the $a\mu$ and the $q_i$ in $aG_K(q_1, \ldots, q_n; p_1, \ldots, p_n)$ as replaced by their values in terms of the $p_K$.)

The integrals $\psi(p_1, \ldots, p_n)$ of $DKdp_K$ and $\phi(q_1, \ldots, q_n)$ of $Cdqi$ are, of course, related as the $\psi$ and $\phi$ of our theorem on extremes. If (12) should happen to be difficult to integrate formally, while (15) is comparatively simple, we can first find $\psi$ and then use the transformation between the $q_i$ and the $p_K$ defined by (13) to compute $\phi$. (15) might be simpler to integrate than (12) because although the integrating factor is the same for both, that is, $\gamma$, the expression for $\gamma$ in terms of the $p_K$ may be simpler than the expression for $\gamma$ in terms of the $q_i$. "Almost any" transformation can be used to introduce a new set of variables $p_K$ in terms of the original variables $q_i$ and then $C_i q_{iK} dp_K$ will be integrable if $Cdqi$ is; the point of the present section is that if functions $aG$ satisfying our assumptions can be found, the correlated total differential $DKdp_K$ can be obtained without having to go back to $Cdqi$ and compute the partial derivatives of the $q_i$ with respect to the $p_K$ (assuming that equations (13) have been solved, we simply use the relation $DK = -a\mu aG_K$).

Bibliography