SOME NONCLASSICAL PROBLEMS IN THE CALCULUS OF VARIATIONS

RICHARD BELLMAN, IRVING GLICKSBERG, AND OLIVER GROSS

1. Introduction. In this paper we shall discuss a number of variational problems of nonclassical variety. They may be considered to arise in the following way.

We have a system of linear inhomogeneous equations of the form

$$(1) \qquad \frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}x_j + f_i(t), \qquad i = 1, 2, \cdots, n, \qquad x_i(0) = c_i,$$

where we may for the sake of simplicity assume the coefficients, a_{ij}, to be constant. The general problem is now that of choosing the "forcing" functions, $f_i(t)$, so as to minimize or maximize some functional of the solutions, $x_i(t)$, subject to constraints of various types.

The simplest case, and readily amenable to classical techniques, occurs when we take the functional to be quadratic, in the form

$$(2) \qquad J(f) = \int_0^T Q(x(t) - g(t))dt,$$

where $Q(u) = \sum_{i,j=1}^{n} b_{ij}u_iu_j$ is a quadratic form in the components of u, and take the constraints on f_i to be of the form

$$(3) \qquad \int_0^T f_i^2 dt \leqq a_i, \qquad i = 1, 2, \cdots, n.$$

This has been treated in [3], in the even more general setting where

$$(4) \qquad x_i(t) = c_i + \int_0^T \left[\sum_{i=1}^{n} K_i(s, t)f_i(s) \right]ds,$$

which includes the case of differential, differential-difference, and more general functional equations.

In [1; 2; 6] we have discussed some problems of this general type when the functionals are all linear. These arise in connection with "bottleneck" problems. Also in [3; 4], we solved some problems involving various combinations of constraints of quadratic and linear form. The technique used there relied heavily on the Neyman-

Pearson lemma, while in $[1; 2; 6]$ we used the functional equation approach of the theory of dynamic programming.

In $[5]$, we considered the problem of determining the $f_i(t)$, subject to $|f_i(t)| \leq 1$, so as to minimize the time required to reduce all the $x_i(t)$ simultaneously to zero. This is the "bang-bang" control problem.

In this paper, continuing the same general line of research, we wish to consider a class of variational problems by conversion into equivalent game-theoretic problems and subsequent application of various extensions of the fundamental min-max theorem of von Neumann in the theory of games.

In order to illustrate the principle, we shall consider two problems connected with the one-dimensional version of (1),

$$(5) \qquad \frac{dx}{dt} = -x + f, \qquad x(0) = 1.$$

The first is that of choosing f subject to

$$(6) \qquad 0 \leq f \leq M, \qquad \int_0^T f\, dt \leq a < MT,$$

so as to minimize

$$(7) \qquad J_1(f) = \int_0^T |1 - x(t)|\, dt,$$

and the second is that of minimizing

$$(8) \qquad J_2(f) = \operatorname*{Max}_{0 \leq t \leq T} |1 - x(t)|$$

subject to the same constraints.

A third problem, which requires a much more detailed discussion, is that of maximizing

$$(9) \qquad J(f, g) = \int_0^T \operatorname{Min}(x, y)\, dt$$

over all f and g subject to the constraints

$$(10) \qquad \begin{aligned} \frac{dx}{dt} &= -ax + f, \qquad x(0) = c_1 > 0, \\[2mm] \frac{dy}{dt} &= -by + g, \qquad y(0) = c_2 > 0. \end{aligned}$$

$$\text{(a) } 0 \leq f \leq 1, \qquad \int_0^T f dt \leq \alpha,$$

(11)

$$\text{(b) } 0 \leq g \leq 1, \qquad \int_0^T g dt \leq \beta.$$

Since the solution requires the consideration of a large number of cases, and is quite long and complicated, we shall not present it here.

At first sight, it might seem that the conversion of a maximization problem into minimum-maximorum or maximum-minimorum problem would hardly lead to a simplification.

As we will see below, however, the new problems will possess two features of importance. In the first place, we can choose the new functionals to be linear, and thus eliminate much of the unpleasantness associated with $|1-x|$ or Min (x, y). In the second place, due to the fact that we are in a game situation where min-max $=$max-min, we can test a conjectured solution quite easily and thus avoid a detailed investigation of local maximum versus global maximum.

2. **The basic device.** To simplify the functionals $|1-x|$ and Min (x, y) we use the simple identities

$$(1) \qquad\qquad |1 - x| = \underset{-1 \leq \phi \leq 1}{\text{Max }} \phi(1 - x)$$

and

$$(2) \qquad\qquad \text{Min } (x, y) = \underset{0 \leq \phi \leq 1}{\text{Min }} (\phi x + (1 - \phi)y).$$

Hence, with ϕ a measurable function,

$$(3) \qquad \underset{f}{\text{Min }} \int_0^T |1 - x(t)| \, dt = \underset{f}{\text{Min }} \underset{-1 \leq \phi \leq 1}{\text{Max }} \int_0^T \phi(1 - x) dt,$$

$$\underset{f,g}{\text{Max }} \int_0^T \text{Min } (x, y) dt = \underset{f,g}{\text{Max }} \underset{0 \leq \phi \leq 1}{\text{Min }} \int_0^T (\phi x + (1 - \phi)y) dt.$$

Using the min-max theorems which we shall discuss in the next section, we have

$$(4) \qquad \underset{f}{\text{Min }} \underset{-1 \leq \phi \leq 1}{\text{Max }} \int_0^T \phi(1 - x) dt = \underset{-1 \leq \phi \leq 1}{\text{Max }} \underset{f}{\text{Min }} \int_0^T \phi(1 - x) dt$$

and

(5) $\displaystyle \underset{f,\theta}{\text{Max}}\ \underset{0 \leqq \phi \leqq 1}{\text{Min}} \int_0^T (\phi x + (1 - \phi)y)dt$

$$= \underset{0 \leqq \phi \leqq 1}{\text{Min}}\ \underset{f,\theta}{\text{Max}} \int_0^T (\phi x + (1 - \phi)y)dt.$$

In each case, the fact that max-min equals min-max means that some pair (f, ϕ) constitute a "saddle point." This enables their properties to be determined in a much easier fashion.

3. **Min-max theorems.** For our purpose the simplest formulation of the min-max relation we shall use is a weakened form of those given by Ky Fan [7] and H. Nikaidô [8]. Their results yield the fact that if X and Y are compact convex subsets of a linear space and M is a real valued function on $X \times Y$, concave on X, convex on Y, and separately continuous, then

(1) $\displaystyle \underset{y}{\text{Min}}\ \underset{x}{\text{Max}}\ M(x, y) = \underset{x}{\text{Max}}\ \underset{y}{\text{Min}}\ M(x, y);$

hence there is a saddle point (x_0, y_0) for which

(2) $M(x_0, y) \geqq M(x_0, y_0) \geqq M(x, y_0), \quad x \in X, y \in Y.$

In our particular cases (where ϕ and f, or (f, g) and ϕ play the roles of x and y), the set of all measurable f for which $0 \leqq f \leqq M$, $\int_0^T f dt \leqq a$, or the set of all measurable ϕ satisfying $0 \leqq \phi \leqq 1$ (or $|\phi| \leqq 1$) may be taken as subsets of $L_\infty(0, T)$, the adjoint of $L_1(0, T)$; as such they are compact in the weak* topology. Since

(3) $\displaystyle \int_0^T \phi(1 - x)dt, \qquad \int_0^T (\phi x + (1 - \phi)y)dt$

are both separately continuous and satisfy the convexity conditions, as is easily seen, we have in both cases a saddle point.

4. **Min $\int_0^T |1 - x(t)| dt$.** Let us now consider in detail the problem of minimizing

(1) $\displaystyle J(f) = \int_0^T |1 - x(t)| dt,$

where x is the absolutely continuous solution of

(2) $dx/dt = -x + f, \qquad x(0) = 1,$

and f belongs to the class of functions, F, satisfying the constraints

$$\text{(a)} \quad 0 \leqq f \leqq M > 1,$$

(3)

$$\text{(b)} \quad \int_0^T f dt \leqq \alpha < T.$$

Let Φ be the set of all measurable ϕ for which $-1 \leqq \phi \leqq 1$. Then

$$\text{(4)} \qquad \underset{f \in F}{\text{Min}} \int_0^1 |1 - x(t)| \, dt = \underset{f \in F}{\text{Min}} \underset{\phi \in \Phi}{\text{Max}} \int_0^T (1 - x(t)) \phi(t) dt.$$

Using the min-max theorem discussed above we know that

$$\text{(5)} \quad \underset{f \in F}{\text{Min}} \underset{\phi \in \Phi}{\text{Max}} \int_0^T (1 - x(t)) \phi(t) dt = \underset{\phi \in \Phi}{\text{Max}} \underset{f \in F}{\text{Min}} \int_0^T (1 - x(t)) \phi(t) dt.$$

Consequently a saddle point of the continuous game with payoff

$$\text{(6)} \qquad M(\phi, f) = \int_0^T (1 - x(t)) \phi(t) dt, \qquad \phi \in \Phi, f \in F,$$

will solve our original minimization problem.

For such a saddle point (ϕ_0, f_0), which we know exists, we have, for almost all t:

(7)
$$x_0(t) < 1 \text{ implies } \phi_0(t) = 1,$$
$$x_0(t) > 1 \text{ implies } \phi_0(t) = -1$$

since ϕ_0 maximizes $M(f_0, \phi)$.

Since

$$M(\phi_0, f) = k(\phi_0) - \int_0^T \phi_0(t) e^{-t} \left[\int_0^T e^s f(s) ds \right] dt$$

(8)

$$= k(\phi_0) - \int_0^T f(s) e^s \left\{ \int_s^T e^{-t} \phi(t) dt \right\} ds,$$

where $k(\phi_0)$ is independent of f, f_0 maximizes the inner product

$$\text{(9)} \qquad (f, K) = \int_0^T f(s) K(s) ds,$$

where

$$\text{(10)} \qquad K(s) = e^s \int_s^T e^{-t} \phi_0(t) dt.$$

From this it follows (as in [3]) that there is a unique λ for which, for almost all t

$$(11) \quad \begin{aligned} K(t) > \lambda \text{ implies } f_0(t) &= M, \\ K(t) < \lambda \text{ implies } f_0(t) &= 0. \end{aligned}$$

On any component, I, of the open set $\{t : K(t) > \lambda\}$ we have $f_0 = M > 1$, almost everywhere. Since

$$(12) \qquad (e^t x_0(t))' = e^t x_0'(t) + e^t x_0(t) = f e^t \leqq M e^t$$

whence, integrating between 0 and t

$$(13) \qquad e^t x_0(t) - 1 \leqq M e^t - M,$$

we have $x_0(t) < M$ for all t. Hence $x_0'(t) = -x_0(t) + M > 0$ almost everywhere on I, and $x_0(t)$ is strictly increasing. If $x_0(t) < 1$ for some t in I, then $[0, t] \subset I$, for if $t_0 > 0$ is the left endpoint of I, we have $x_0(s) < 1$ and thus $\phi_0(s) = 1$ for $t_0 \leqq s \leqq t$. Hence

$$(14) \quad K'(s) = K(s) - \phi(s) = K(s) - 1 < e^s \int_s^\infty e^{-t} dt - 1 = 0,$$

almost everywhere on (t_0, t), and $K(t_0) > \lambda$. But this implies that a neighborhood $(t_0 - \epsilon, t_0 + \epsilon)$ of t_0 is in I, a contradiction, unless $t_0 = 0$.

Since x_0 is increasing on I and $x_0(0) = 1$, we must have $x_0(t) > 1$, contradicting our assumption. Hence $x_0 \geqq 1$ on I.

Thus, if I is nonvoid, there is a $t \in I$ for which $x_0(t) > 1$. In this case $(t, T) \in I$. For if t_0 is the upper endpoint of I, $t_0 > t$, then $x_0(s) > 1$ and $\phi_0(s) = -1$ almost everywhere on (t, t_0) so that $K'(s) = K(s) + 1 \geqq 1 - e^s \int_s^T e^{-t} dt > 1 - e^s \int_s^\infty e^{-t} dt = 0$ there. Hence $K(t_0) > \lambda$, whence as before $t_0 = T$. Thus $x_0(T) > 1$.

If we set

$$(15) \quad \begin{aligned} f(t) &= f_0(t), & 0 \leqq t \leqq T - \epsilon, \\ f(t) &= 0, & t \geqq T - \epsilon, \end{aligned}$$

for some small positive ϵ, we obviously have

$$(16) \qquad \int_0^T |1 - x(t)| \, dt < \int_0^T |1 - x_0(t)| \, dt,$$

contradicting the fact that f_0 minimizes the integral.

We must conclude therefore that the set $\{t : K(t) > \lambda\}$ is void, and that $K \leqq \lambda$. On the set where $K = \lambda$, or

$$(17) \qquad \lambda e^{-s} = \int_s^T e^{-t} \phi_0(t) dt,$$

we must have $\phi_0 = \lambda$, a. e. Since $|K| < 1$, we see that $|\lambda| < 1$. Hence

$x_0 = 1$ almost everywhere on the set where $K = \lambda$, by (7).

Since $f_0 = 0$ on the open set $\{t : K(t) < \lambda\}$, x_0 is decreasing on its components, so that obviously the set $\{t : K(t) = \lambda\}$ in an initial subinterval of $[0, T]$. Clearly $f_0 = 1$ there. Hence f_0 is the characteristic function of the interval $[0, a]$.

5. Max $|1 - x(t)|$. Let us consider next the problem of minimizing

$$(1) \qquad\qquad J(f) = \underset{0 \leqq t \leqq T}{\mathrm{Max}} \; |1 - x(t)|$$

where as before

$$x(t) = e^{-t} + e^{-t} \int_0^t e^s f(s) ds, \qquad\qquad 0 \leqq f \leqq M \; (M > 1)$$

and $\int_0^T f dt \leqq \alpha < T$. Since

$$(2) \qquad\qquad \underset{0 \leqq t \leqq T}{\mathrm{Max}} \; |1 - x(t)| \; = \mathrm{Max}_\mu \int_0^T (1 - x(t)) d\mu(t)$$

where μ varies over all signed measures of total variation at most 1 we have a corresponding game with payoff

$$(3) \qquad\qquad N(f, \mu) = \int_0^T (1 - x(t)) d\mu(t).$$

Here we again have a compact set of f and μ (the latter set taken in the weak* topology of the adjoint of $C(0, 1)$). M is separately continuous and appropriately convex and concave so that we are assured that a saddle point (f_0, μ_0) exists. This f_0 will of course solve our problem since, by (2),

$$(4) \qquad\qquad N(f_0, \mu_0) = \underset{f}{\mathrm{Min}} \; \underset{\mu}{\mathrm{Max}} \; N(f, \mu)$$

$$= \underset{f}{\mathrm{Min}} \; \underset{t}{\mathrm{Max}} \; |1 - x(t)|.$$

Although one could now proceed, as in the previous section, with a tedious deduction of f_0, there is an obvious guess: that

$$(5) \qquad\qquad f_0(t) = \begin{cases} 0, & t \leqq t_0, \\ x_0(t_0), & t_0 < t \leqq T, \end{cases}$$

where $x_0(t_0)(T - t_0) = \alpha = e^{-t_0}(T - t_0)$; thus x_0 decreases from 1 to its minimum $x(t_0)$ which it assumes on $[t_0, T]$. If this guess is correct we

must obviously have $d\mu_0 = 0$ for $t < t_0$, $d\mu_0 \geqq 0$, and μ_0's total increase 1, in view of (2). Moreover, since

$$
\begin{aligned}
(6) \quad \int_0^T (1 - x(t))d\mu_0(t) &= k(\mu_0) - \int_0^T e^{-t} \int_0^t e^s f(s) ds d\mu_0(t) \\
&= k(\mu_0) - \int_0^T f(s) e^s \int_s^T e^{-t} d\mu_0(t) ds
\end{aligned}
$$

is minimized by $f = f_0$ and $M > 1 \geqq f_0$ we must have, by the Neyman-Pearson lemma, a $\lambda \geqq 0$ for which for $s \geqq t_0$, $\lambda = e^s \int_s^T e^{-t} d\mu_0(t)$, for $s < t_0$, $\lambda \leqq e^s \int_s^T e^{-t} d\mu_0(t)$. Both conditions are satisfied by setting

$$
\begin{aligned}
(7) \quad \mu_0(t) &= 0, & t &\leqq t_0, \\
\mu_0(t) &= \lambda(t - t_0), & t_0 &< t < T, \\
\mu_0(T) &= \lambda(T - t_0) + \lambda = 1,
\end{aligned}
$$

since then $e^s \int_s^T e^{-t} d\mu_0 = e^s \int_s^T e^{-t} du_0 + e^s e^{-T} \lambda = \lambda e^s (e^{-s} - e^{-T}) + e^{s-T} \lambda = \lambda$ for $s > t_0$, while for $s < t_0$, $e^s \int_s^T e^{-t} d\mu_0 = e^{s-t_0} e^{t_0} \int_{t_0}^T e^{-t} d\mu_0 = e^{s-t_0} \lambda < \lambda$. Thus it is clear that $N(f_0, \mu_0) \leqq N(f, \mu_0)$ while since μ_0 has its total increase in the set where $1 - x_0(t) = \text{Max} \, |1 - x(t)|$, $N(f_0, \mu_0) = \text{Max} \, |1 - x_0(t)| \geqq N(f_0, \mu)$, which completes our verification.

BIBLIOGRAPHY

1. R. Bellman, *Bottleneck problems and dynamic programming*, Proc. Nat. Acad· Sci. U.S.A. vol. 39 (1953) pp. 947–951.

2. ———, *Bottleneck problems, functional equations and dynamic programming*, Econometrica vol. 23 (1955) pp. 78–87.

3. R. Bellman, I. Glicksberg, and O. Gross, *On some variational problems occurring in the theory of dynamic programming*, Proc. Nat. Acad. Sci. U.S.A. vol. 39 (1953) pp. 298–301.

4. ———. *On some variational problems occurring in the theory of dynamic programming*, Rend. Circ. Mat. Palermo (2) vol. 3 (1954) pp. 363–397.

5. ———, *On the "bang-bang" control problem*, Quarterly of Applied Mathematics (to appear).

6. R. Bellman and S. Lehman, *On the continuous gold-mining equation*, Proc. Nat. Acad. Sci. U.S.A. vol. 40 (1954) pp. 115–119.

7. Ky Fan, *Minimax theorems*, National Bureau of Standards Report, December 10, 1952.

8. H. Nikaidô, *On von Neumann's minimax theorem*, Pacific Journal of Mathematics vol. 4 (1954) pp. 65–72.

RAND CORPORATION