

Direct Transformation of Variational Problems into Cauchy Systems—II. Scalar-Semi-Quadratic Case*

by JAMES HESS, HARRIET KAGIWADA and ROBERT KALABA

*Department of Economics, University of Southern California,
Los Angeles, CA 90007 and HFS Associates 3117 Malcolm Ave.,
Los Angeles, CA 90034*

ABSTRACT: *This series of papers addresses three interrelated problems: the solution of a variational problem, the solution of integral equations, and the solution of an initial valued system of integrodifferential equations. It will be shown that a large class of variational problems requires the solution of a nonlinear integral equation. It has also been shown that the solution of a nonlinear integral equation is identical to the solution of a Cauchy system. In this paper, we by-pass the nonlinear integral equations and show that the minimization problems directly imply a solution of the Cauchy system. This second paper in the series looks at semi-quadratic functional and scalar functions.*

1. Introduction

Many optimization problems result in variational problems of finding a function $z(t)$, $0 \leq t \leq 1$, that minimizes the functional

$$W[z] = \lambda \int_0^1 \int_0^1 z(s)k(t, s)z(t) dt ds + 2 \int_0^1 F(z(t), t) dt.$$

By standard variational techniques it can be shown that the optimal function, $u(t)$, satisfies a nonlinear Fredholm integral equation [Ref. (1)]. Recent work in the study of integral equations [Ref. (2)] has shown that solutions of nonlinear integral equations are equivalent to solutions of particular initial valued systems of integrodifferential equations (Cauchy Systems). It appears that the three problems—variational, integral equations, and Cauchy Systems—are equivalent to each other. An important missing link in the analysis has been the demonstration that the variational problem leads directly to a Cauchy System without passing through an integral equation. These papers provide that link.

* Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. 77-3383.

II. Derivation

Suppose we desire to find a scalar function $z(t)$, $0 \leq t \leq 1$, which minimizes the semi-quadratic functional

$$W[z, \lambda] = \lambda \int_0^1 \int_0^1 z(s)k(t, s)z(t) ds dt + 2 \int_0^1 F(z(t), t) dt, \quad (1)$$

where $k(t, s)$ is a symmetric, positive definite kernel, $F(z, t)$ is a convex, twice differentiable function in z , and λ is a sufficiently small scalar parameter. This class of problems is semi-quadratic because the first term is a quadratic functional in z . This is more general than it might look at first glance, for the choice variable may actually be a function $x(t)$ and $z(t)$ might be a composite function,

$$z(t) = h(x(t)) \quad (2)$$

where h is a convex, differentiable function. By simple change of variables this more general problem may be reduced to Eq. (1). The standard variational approach to this problem results in a nonlinear integral equation which the optimal function must satisfy.

Proposition 1. The function $u(t)$ which minimizes $W[z, \lambda]$ must satisfy the integral equation

$$F_1(u(t), t) + \lambda \int_0^1 k(t, s)u(s) ds = 0, \quad 0 \leq t \leq 1. \quad (3)$$

Throughout the paper, the partial derivative of F with respect to its i th argument is denoted by F_i and the derivative of $u(t, \lambda)$ with respect to λ is denoted by $u_\lambda(t, \lambda)$.

The proposition is established as follows. The arbitrary admissible functions may be written as

$$z(t) = u(t) + \epsilon \eta(t), \quad (4)$$

where ϵ is an arbitrary scalar, and $\eta(t)$ is an arbitrary function. Take a Taylor series expansion of $W[z, \lambda]$ in ϵ , ignoring all terms in ϵ^2 or higher.

This results in

$$W[z, \lambda] = W[u, \lambda] + \epsilon C[u, \eta, \lambda] \quad (5)$$

where

$$C[u, \eta, \lambda] = \lambda \int_0^1 \int_0^1 k(t, s)u(s)\eta(t) ds dt + \lambda \int_0^1 \int_0^1 u(t)k(t, s)\eta(s) ds dt + 2 \int_0^1 F_1(u(t), t)\eta(t) dt. \quad (6)$$

Optimality requires

$$\epsilon C[u, \eta, \lambda] \geq 0 \quad (7)$$

for all ϵ, η . Since ϵ has arbitrary sign, this implies

$$C[u, \eta, \lambda] = 0. \quad (8)$$

Using the symmetry of $k(t, s)$, variables of integration may be relabeled so that Eq. (8) is expressed as

$$\lambda \int_0^1 \int_0^1 k(t, s)u(s)\eta(t) ds dt + \int_0^1 F_1(u(t), t)\eta(t) dt = 0 \quad (9)$$

for all arbitrary $\eta(t)$. Applying the fundamental lemma of the calculus of variations to Eq. (9) results in the desired integral equation. This completes the proof.

Another approach to the minimization problem is to ask how the optimal solution changes as λ varies. This is referred to as *parameter imbedding*. The basic idea of parameter imbedding is to convert the nonlinear integral equation into a system of initial valued integrodifferential equations. The basic result given in Ref. (2) is:

Proposition 2. The function $u(t) = u(t, \lambda)$ which satisfies the nonlinear integral Eq. (3) is the solution to the following initial valued integrodifferential equations, involving a resolvent kernel $K(t, s, \lambda)$, and conversely:

$$u_\lambda(t, \lambda) + \alpha(t, \lambda) + \lambda \int_0^1 K(t, s, \lambda)\alpha(s, \lambda) ds = 0 \quad (10)$$

$$K_\lambda(t, s, \lambda) + \delta(t, s, \lambda) + \int_0^1 K(t, s', \lambda)\delta(s', s, \lambda) ds' = 0 \quad (11)$$

$$F_1(u, (t, 0), t) = 0 \quad (12)$$

$$K(t, s, 0) + k(t, s)/F_{11}(u(t, 0), t) = 0 \quad (13)$$

where

$$\alpha(t, \lambda) = \int_0^1 k(t, s)u(s, \lambda) ds/F_{11}(u(t, \lambda), t) \quad (14)$$

$$\delta(t, s, \lambda) = \gamma(t, s, \lambda) + \int_0^1 k(t, s')K(s', s, \lambda)/F_{11}(u(t, \lambda), t) ds'$$

$$+ \lambda \int_0^1 \gamma(t, s', \lambda)K(s', s, \lambda) ds' \quad (15)$$

$$\gamma(t, s, \lambda) = \frac{\partial}{\partial \lambda}(k(t, s)/F_{11}(u(t, \lambda), t)), \quad (16)$$

for $0 \leq t \leq 1$, $0 \leq s \leq 1$.

The Cauchy System (10)–(16) has been very useful for the computation of solutions of nonlinear integral equations, whether or not they arise from a variational problem. Well-known techniques, such as the Runge–Kutta or Adams–Moulton methods, together with the methods of lines, are readily available. [See Ref. (3) for one such example.] In addition, there are many cases where the parameter λ has an interesting physical interpretation and thus the

Cauchy System provides all the equations needed to study the sensitivity of the solution to changes in important parameters

It will now be shown that the Cauchy System (10)–(16) may be derived directly from the semi-quadratic minimization problem without ever writing down the nonlinear integral Eq. (3). This shows that the solution of the variational problem could have proceeded even if the integral equation had never been discovered. The Cauchy System is perfectly adequate for describing the optimal function $u(t, \lambda)$.

Proposition 3. The function $u(t, \lambda)$ which minimizes $W[u, \lambda]$ must satisfy the Cauchy System (10)–(16).

The proposition is established as follows. Let $u(t, \lambda)$ be the solution of the minimization problem for parameter value λ , and let $u(t, \lambda + d\lambda)$ be the solution for $\lambda + d\lambda$. Admissible solutions to the variational problem may be expressed as

$$z(t, \lambda) = u(t, \lambda) + \epsilon \eta(t), \tag{17}$$

$$x(t, \lambda + d\lambda) = u(t, \lambda + d\lambda) + \sigma \rho(t), \tag{18}$$

where ϵ and σ are arbitrary scalars and where η and ρ are arbitrary functions. Approximate $W[z, \lambda]$ and $W[x, \lambda + d\lambda]$ by a Taylor series in the first argument. When ϵ and σ are suitably small the terms in ϵ^2 and σ^2 or higher may be ignored. The resulting approximations are

$$W[z, \lambda] \cong W[u, \lambda] + \epsilon C[u, \eta, \lambda] + \text{higher order terms} \tag{19}$$

$$W[x, \lambda + d\lambda] \cong W[u, \lambda + d\lambda] + \sigma C[u, \rho, \lambda + d\lambda] + \text{higher order terms} \tag{20}$$

where

$$\begin{aligned} C[u, \eta, \lambda] &= 2\lambda \int_0^1 \int_0^1 k(t, s) u(s, \lambda) \eta(t) ds dt \\ &+ 2 \int_0^1 F_1(u(t, \lambda), t) \eta(t) dt \end{aligned} \tag{21}$$

$$\begin{aligned} C[u, \rho, \lambda + d\lambda] &= 2(\lambda + d\lambda) \int_0^1 \int_0^1 k(t, s) u(s, \lambda + d\lambda) \rho(t) ds dt \\ &+ 2 \int_0^1 F_1(u(t, \lambda + d\lambda), t) \rho(t) dt. \end{aligned} \tag{22}$$

If $u(t, \lambda)$ and $u(t, \lambda + d\lambda)$ are optimal, then it must be true that

$$\epsilon C[u, \eta, \lambda] \geq 0 \tag{23}$$

$$\sigma C[u, \rho, \lambda + d\lambda] \geq 0 \tag{24}$$

for all $\epsilon, \sigma, \eta, \rho$. Since these are arbitrary, select $\epsilon = -\sigma$ and $\eta = \rho$ so that (23) and (24) are

$$-\sigma C[u, \rho, \lambda] \geq 0, \tag{25}$$

$$\sigma C[u, \rho, \lambda + d\lambda] \geq 0. \tag{26}$$

Direct Transformation of Variational Problems

Add these two inequalities to get

$$\sigma C[u, \rho, \lambda + d\lambda] - C[u, \rho, \lambda] \geq 0. \tag{27}$$

Since σ has arbitrary sign, this implies that

$$C[u, \rho, \lambda + d\lambda] - C[u, \rho, \lambda] = 0, \tag{28}$$

for all arbitrary ρ . Expand $C[u, \rho, \lambda + d\lambda]$ in a Taylor series in $\lambda + d\lambda$; ignoring terms with $(d\lambda)^2$ or higher this gives

$$\begin{aligned} C[u, \rho, \lambda + d\lambda] &\cong C[u, \rho, \lambda] + d\lambda \{ 2 \int_0^1 \int_0^1 k(t, s) u(s, \lambda) \rho(t) ds dt \\ &+ 2\lambda \int_0^1 \int_0^1 k(t, s) u_s(s, \lambda) \rho(t) ds dt \\ &+ 2 \int_0^1 F_{11}(u(t, \lambda), t) u_t(t, \lambda) \rho(t) dt \}. \end{aligned} \tag{29}$$

Hence for $d\lambda$ sufficiently small, Eqs. (28) and (29) imply that the term in brackets in Eq. (29) must be zero for all arbitrary $\rho(t)$. Apply the fundamental lemma of the calculus of variations and we have

$$\begin{aligned} F_{11}(u(t, \lambda), t) u_t(t, \lambda) &+ \int_0^1 k(t, s) u(s, \lambda) ds \\ &+ \lambda \int_0^1 k(t, s) u_s(s, \lambda) ds = 0. \end{aligned} \tag{30}$$

Divide Eq. (30) by $F_{11}(u(t, \lambda), t)$ and we have the linear Fredholm integral equation in the unknown function $u_t(t, \lambda)$

$$u_t(t, \lambda) + \alpha(t, \lambda) + \lambda \int_0^1 m(t, s, \lambda) u_s(s, \lambda) ds = 0, \tag{31}$$

where $\alpha(t, \lambda)$ is defined in Eq. (14) and $m(t, s, \lambda)$ is defined by

$$m(t, s, \lambda) = k(t, s) / F_{11}(u(t, \lambda), t). \tag{32}$$

A linear Fredholm integral equation with kernel $m(t, s, \lambda)$ has a solution that may be expressed using a resolvent kernel $K(t, s, \lambda)$ as follows:

$$u_t(t, \lambda) + \alpha(t, \lambda) + \lambda \int_0^1 K(t, s, \lambda) \alpha(s, \lambda) ds = 0. \tag{33}$$

This is Eq. (10) of the Cauchy System. The resolvent kernel must satisfy a related linear Fredholm integral equation [see Ref. (2)],

$$K(t, s, \lambda) + m(t, s, \lambda) + \lambda \int_0^1 m(t, s', \lambda) K(s', s, \lambda) ds' = 0. \tag{34}$$

James Hess, Harriet Kagiwada and Robert Kalaba

Differentiate (34) with respect to λ to get

$$\begin{aligned} K_\lambda(t, s, \lambda) + \gamma(t, s, \lambda) + \int_0^1 m(t, s', \lambda) K(s', s, \lambda) ds' \\ + \lambda \int_0^1 \gamma(t, s', \lambda) K(s', s, \lambda) ds' \\ + \lambda \int_0^1 m(t, s', \lambda) K_\lambda(s', s, \lambda) ds' = 0, \end{aligned} \quad (35)$$

where $\gamma(t, s, \lambda)$ is defined in Eq. (16). Since Eq. (35) is also a linear Fredholm integral equation with a kernel $m(t, s', \lambda)$, its solution may be expressed using the same resolvent kernel,

$$K_\lambda(t, s, \lambda) + \delta(t, s, \lambda) + \int_0^1 K(t, s', \lambda) \delta(s', s, \lambda) ds' = 0 \quad (36)$$

where $\delta(t, s, \lambda)$ is defined in Eq. (15). This is Eq. (11) of the Cauchy system. When $\lambda = 0$ in the original minimization problem, $W[u(t, 0), 0]$ is minimized when $F(u(t, 0), t)$ is minimized for each t . This implies that

$$F_t(u(t, 0), t) = 0 \quad (37)$$

which is the initial condition (12). The initial condition (13) follows from Eq. (34) by setting $\lambda = 0$. This completes the proof.

III. Discussion

The objective of this paper has been to reduce the semi-quadratic variational problem in Eq. (1) to the Cauchy System in relations (10)–(16). In particular we have been able to do this without making any use of the Euler equation, which takes the form of the nonlinear integral equation (3).

The earlier paper in this series, Ref. (4), demonstrated that the reduction could be performed in the quadratic case. In this paper, the general technique of direct reduction is made more explicit. The objective functional is approximated first in a Taylor series in the choice function and then in a Taylor series in the parameter λ . It would have been possible to derive the nonlinear integral equation along the way but this was not done in order to arrive at the sensitivity Eqs. (10)–(16). This general technique may be applied to problems which are not semi-quadratic and which have a vector of choice functions. This approach may be used to develop new equations for such variational problems as the team decision problem, [Ref. (5)], or the simplest problem in the calculus of variations.

References

- (1) R. Courant and D. Hilbert, "Methods of Mathematical Physics", John Wiley Interscience, New York, 1953.

- (2) H. Kagiwada and R. Kalaba, "Integral Equations Via Imbedding Methods", Addison-Wesley, Reading, Mass., 1974.
- (3) J. Casti and R. Kalaba, "Imbedding Methods in Applied Mathematics", Addison-Wesley, Reading, Mass., 1973.
- (4) J. Hess and R. Kalaba, "Direct transformation of variational problems into Cauchy systems—I. Scalar-quadratic case", *J. Optimization Theory and Applications*, Vol. 24, No. 4, 1978.
- (5) J. Marschak and R. Radner, "Economic Theory of Teams", Yale University Press, New Haven, Conn., 1972.