Embedded Construction of Adjoint Equations for Optimization using Continuation

Mingwu Li*, Cole Anderson* and Harry Dankowicz*
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Illinois, United States

Summary. We present a novel, object-oriented framework for staged construction of adjoint equations for optimization problems with algebraic and differential constraints, and illustrate its application to the problem of maximizing the dynamic range of a microbeam oscillator using successive stages of numerical continuation. The new framework is encoded in production-ready toolboxes compatible with the COCO continuation package that support a straightforward hierarchical construction of composite optimization problems.

Introduction

Numerical continuation is a well-established tool for analyzing nonlinear dynamical systems that has made its way into several successful software implementations. Among these, the MATLAB-based computational continuation core COCO [1] uses an embeddable construction philosophy where larger problems are assembled from smaller ones. The present work demonstrates the integration of a technique for constrained nonlinear optimization in the COCO framework, by relying on the *successive continuation algorithm* developed in [2], that augments an original continuation problem with a set of adjoint equations and locates a local extremum in several consecutive stages. The result is a simplification of the construction of constrained nonlinear optimization problems that generalizes the benefits of COCO for composite continuation problems to composite optimization problems defined, for example, in terms of the properties of isolas of periodic orbits.

Example problem statement and adjoint formulation

Consider the following optimization problem with algebraic, differential and integral constraints.

$$\begin{aligned} \text{minimize:} \quad & \Phi(x(0), x(T), T, p) + \textstyle\int_0^T g(x,p)dt \\ \text{subject to:} \quad & \dot{x} - f(x,p) = 0, \;\; \Psi(x(0), x(T), T, p) = 0, \;\; \textstyle\int_0^T h(x,p)dt = 0 \end{aligned}$$

in terms of an unknown trajectory $x(t)$, interval duration T, and problem parameters p. We use the method of Lagrange multipliers to obtain the optimality system. After rescaling time by the transformation $\{t = T\tau \mid \tau \in [0,1]\}$, the following Lagrangian, L, and first-order variational optimality condition, $\delta L = 0$, yield the expanded system of equations given in Table 1, where λ_i's and η_i's are corresponding Lagrange multipliers.

$$L = \mu_1 + \int_0^1 \lambda_1^T(\tau)\left(\dot{x} - T\,f\right)d\tau + \lambda_2^T\,\Psi + \lambda_3 \int_0^1 h d\tau + \eta_1\left(\Phi + T\int_0^1 g d\tau - \mu_1\right) + \eta_2^T\left(p - \mu_2\right)$$

Original System		Adjoint System	
Variation	Equation	Variation	Equation
$\delta\lambda_1$	$\dot{x} - T\,f = 0$	δx	$-\dot{\lambda}_1^T - T\lambda_1^T\,f_x + \eta_1\,T\,g_x + \lambda_3 h_x = 0$
$\delta\eta_2$	$p - \mu_2 = 0$	$\delta x\vert_{\tau=0}$	$-\lambda_1^T(0) + \lambda_2^T\Psi_{x(0)} + \eta_1\Phi_{x(0)} = 0$
$\delta\lambda_2$	$\Psi = 0$	$\delta x\vert_{\tau=1}$	$\lambda_1^T(1) + \lambda_2^T\,\Psi_{x(1)} + \eta_1\Phi_{x(1)} = 0$
$\delta\lambda_3$	$\int_0^1 h d\tau = 0$	δT	$-\int_0^1 \lambda_1^T f + \lambda_2^T\Psi_T + \eta_1\Phi_T + \eta_1\int_0^1 g d\tau = 0$
$\delta\eta_1$	$\Phi + T\int_0^1 g d\tau - \mu_1 = 0$	δp	$-T\int_0^1 \lambda_1^T\,f_p d\tau + \lambda_2^T\,\Psi_p + \eta_1\Phi_p + T\eta_1\int_0^1 g_p d\tau + \eta_2^T = 0$
		$\delta\mu_1$	$1 - \eta_1 = 0$
		$\delta\mu_2$	$\eta_2^T = 0$

Table 1: Optimality system obtained from the variational condition $\delta L = 0$.

In particular, consistent with the implementation in COCO, the continuation parameters μ_1 and μ_2 are constrained to equal the cost objective and the problem parameters, respectively. The adjoint equations constitute a necessary conditions for a local extremum of the Lagrangian along the solution manifold associated with the differential, algebraic, and integral constraints. Notably, at such an extremum, we must have $\eta_1 = 1$ and $\eta_2 = 0$.

Continuation algorithm

Our solution approach, inspired by [2], locates an extremum using at least two consecutive applications of a parameter continuation algorithm to the optimality system minus the explicit conditions on η_1 and η_2. To this end, consistent with the COCO implementation, introduce the additional continuation parameters, $\mathrm{d}\mu_1$ and $\mathrm{d}\mu_2$ that equal the variables η_1 and η_2. In the first stage, initialize the λ_i's and η_i's to zero, and allow μ_1, $\mathrm{d}\mu_1$, all but one component of $\mathrm{d}\mu_2$ and the corresponding component of μ_2 to vary, while keeping the remaining components of μ_2 and $\mathrm{d}\mu_2$ fixed, corresponding to continuation along a one-dimensional solution manifold with the λ_i's and η_i's constantly equal to zero. From a geometric fold in μ_1 along this manifold, continue along a secondary branch with nontrivial values of the λ_i's and η_i's until $\mathrm{d}\mu_1 = 1$. Finally, fix $\mathrm{d}\mu_1$, allow all components of μ_2 to vary, and perform continuation until all components of $\mathrm{d}\mu_2$ equal 0.

Embeddable Construction

The adjoint system, minus the explicit constraints on η_1 and η_2, is linear in the adjoint variables and may be written in the more compact form

$$\begin{pmatrix} \lambda_1^T(\tau) & \eta_2^T & \lambda_2^T & \lambda_3 & \eta_1 \end{pmatrix} \cdot J = 0$$

in terms of a linear operator J (or a suitable discretization) resulting from the variation with respect to the continuation variables x, T, and p. In matrix form, the structure of J is given by

$$J = \begin{pmatrix} \dot{(\cdot)} + Tf_x & -(\cdot)|_{\tau=0} & (\cdot)|_{\tau=1} & -\langle(\cdot), f\rangle & -T\langle(\cdot), f_p\rangle \\ 0 & 0 & 0 & 0 & I_q \\ 0 & \Psi_{x(0)} & \Psi_{x(1)} & \Psi_T & \Psi_p \\ -h_x & 0 & 0 & 0 & 0 \\ -Tg_x & \Phi_{x(0)} & \Phi_{x(1)} & \Phi_T + \int_0^1 g d\tau & \Phi_p + T\int_0^1 g_p d\tau \end{pmatrix}$$

where $\lambda_1^T\dot{(\cdot)} = \dot{\lambda}_1^T, \lambda_1^T(\cdot)|_{\tau=0} = \lambda_1^T(0), \lambda_1^T(\cdot)|_{\tau=1} = \lambda_1^T(1), \lambda_1^T < \cdot, f >= \int_0^1 \lambda_1^T f d\tau$, and $\lambda_1^T < \cdot, f >= \int_0^1 \lambda_1^T f d\tau$. Using COCO's problem construction paradigm, we generate the above coefficient matrix of the adjoint formulation simultaneously with the original system of constraints, stage by stage. At each stage, we add rows and columns to the coefficient matrix corresponding to additional constraints and additional adjoint variables, respectively. For example, declaration of the continuation parameter vector μ_2, corresponding to the constraint $p - \mu_2 = 0$, after the initial imposition of the differential constraint results in the addition of the second row of the matrix representation of J.

Composite Problems

The form of the coefficient matrix J depends on the order in which the constraints are imposed. In any construction, each new constraints results in the addition of rows and, possibly, columns to J. This also generalizes to a truly composite constrained optimization problem, for example one in which multiple, initially independent differential constraints are imposed successively and subsequently glued together. As an example, in this presentation, we illustrate the staged construction of a composite optimization problem for maximizing the dynamic range of a nonlinear oscillator of the hardening type as described in [3]. In one implementation that couples a finite-element discretization of the beam dynamics with parameter continuation of periodic orbits, the optimization problem accounts for the dynamics of multiple periodic orbits (each with its own adjoint system), conditions imposed between periodic orbits (both to achieve the near-linear response and to account for equality of parameters between orbits), and the objective function that will be optimized (which itself may relate the periodic orbits). In this case, the matrix representation of J is of the form

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_N \\ X_1 & X_2 & \dots & X_N \\ Y_1 & Y_2 & \dots & Y_N \end{pmatrix}$$

where the J_i's represent coefficient matrices for both the original and adjoint system of the individual periodic orbits, the X_i's are coefficients related to conditions imposed between periodic orbits, and the Y_i's are coefficients related to the objective function.

Conclusion

Single-objective optimization in the presence of algebraic, differential, and integral constraints can be formulated in terms of a sequence of one and multi-dimensional continuation problems. Our implementation in the software package COCO allows for automated embeddable construction of the adjoint equations associated with composite optimization problems, providing direct access to the optimality system for hierarchical problems in terms of coupled finite- and infinite dimensional constraints.

References

[1] Dankowicz H., Schilder F. (2013) Recipes for Continuation. SIAM.

[2] Kernévez, J.P., Doedel, E.J. (1987) "Optimization in bifurcation problems using a continuation method," in *Bifurcation: Analysis, Algorithms, Applications*, eds. Küpper, T., Seydel, R., and Troger, H, Birkhäuser Basel.

[3] Saghafi, M., Dankowicz H., and Lacarbonara F. (2015) "Nonlinear tuning of microresonators for dynamic range enhancement," *Proc. R. Soc. A.* **471(2179)**, art. no. 20140969.