Solving optimal control problems using fast automatic differentiation

E. S. Zasukhina

Computing Center RAS, zasuhina@ccas.ru

Let $W : R^n \times R^r \to R^1$, $\Phi : R^n \times R^r \to R^n$ be twice continuously differentiable mappings. Suppose that $x \in R^n \quad u \in R^r$ satisfy a system of n scalar algebraic constraint equations

$$\Phi(x, u) = 0_n,$$

where 0_n is a zero n-dimensional vector. It is assumed that $\Phi_x^\top(x, u)$ is nonsingular. Then, by the implicit function theorem, there exists a function $x = x(u) \in C^2(u)$ such that $\Phi(x(u), u) = 0_n$ in some vicinity of (x, u). The variable x is called dependent and u is called independent variable or control. Thus the composite function $\Omega(u) = W(x(u), u)$ is twice differentiable. The gradient $d\Omega(u)/du$ is given by the formula

$$d\Omega(u)/du = W_u(x(u), u) + x_u^\top \; W_x(x(u), u).$$

By the implicit function theorem the derivative x_u is determined from the equation

$$\Phi_u + \Phi_x \; x_u = 0_{nr},$$

where 0_{nr} is an $(n \times r)$ zero matrix. Thus to determine the gradient of the function $\Omega(u)$ it is necessary to solve rn linear equations.

According to methodology of fast automatic differentiation (FAD-methodology) [1] the relations for determining the gradient of the function $\Omega(u)$ are of the form

$$L_x(x, u, p) = W_x(x, u) + \Phi_x^\top(x, u)p = 0_n,$$
$$L_u(x, u, p) = d\Omega/du = W_u(x, u) + \Phi_u^\top(x, u)p,$$

where $L(x, u, p) = W(x, u) + p^\top \Phi(x, u)$ is Lagrange function and $p \in R^n$ is a Lagrange multiplier. Thus, the gradient of $\Omega(u)$ can be obtained by solving one linear system of n equations.

In the report the FAD-methodology is extending to the case of determining second order derivatives of the function $\Omega(u)$. Derived formulas for determining second order derivatives contain adjoint variables. These variables can be obtained by solving an associated matrix equation. The number of scalar equations in this equation is a linear function of the dimension of independent variable. The formulas are adapted to multistep processes resulting from discrete approximation of two optimal control problems. The first one is a problem of optimal control of a process governed by system of ordinary differential equations (in general form). The second one is a problem of optimal control of solutions for the one-dimensional unsteady Burgers equation by means of boundary conditions. The first problem was approximated with the use of Runge–Kutta method of arbitrary order. The structure of arising linear systems with adjoint variables is investigated. It appears that this structure is rather simple. Therefore the solution of these systems doesn't cause difficulties. Taking into account special features of the summands in the formula for determining second order derivatives one can build efficient algorithm of these derivatives calculation. This algorithm enables to decrease computing costs and the required memory. This approach was applied to numerical solution of some discrete optimization problems: problems obtained by approximating a concrete optimal control problem with the use of Runge–Kutta method of various orders and a discrete analogue of optimal control problem for the Burgers equation with concrete values of parameters. All problems were solved by Newton and gradient methods. The step in the chosen direction was determined by minimizing a function which interpolated the objective function along this direction by cubic splines. Near the solution Newton method was switched to the standard Newton method. The results of numeric experiments are discussed and analyzed.

References

[1] Yu. G. Evtushenko, *Computation of exact gradients in distributed dynamic systems.*//Optimization methods and software. 1998. V. 9. P. 45-75.