Adjoint based gradient calculation
- advantantages and challenges

Ruben Ringset - student at Department of Engineering Cybernetics, NTNU

Most optimization techniques for nonlinear model predictive control (NMPC) are gradient based approaches. These are often iterative algorithms where derivative information is required at every iteration. Gradient calculation by the adjoint method, also known as the reverse method is based on introducing extra variables into the problem. By letting these variables satisfy some special condition known as the adjoint equations, we can avoid doing certain parts of the gradient calculation. That is, we pay a price by calculating the adjoint equations and we gain something from canceling out other calculations.

Roughly speaking, the adjoint, or the reverse method is favorable when we want to compute sensitivities of a scalar or low dimensional function with respect to many parameters. In the opposite case where we want to compute sensitivities from few parameters to a high dimensional function, forward sensitivity analysis is probably better suited.

When calculating the gradient of the NMPC objective function we are in the position where we want sensitivities from all the control variables with respect to the scalar objective. This motivates us to look into how adjoint sensitivity analysis applies to this problem.

In the case of forward sensitivity analysis for NMPC, a lot of system simulations must be performed in order to first obtain the sensitivities from the control variables to the state variables. These sensitivities are in turn is used to obtain the sensitivities to the outputs by applying the chain rule.

By using the adjoint approach we can omit computation of the sensitivities from the control variables to the state variables by instead solving the adjoints equations. When using this method, the gradient of the objective function can be obtained by only two simulations. Given a sequence of controls, the system equations are first simulated forward in time to obtain the state variables and the output variables. Then we only need to do one simulation backwards in time to obtain the desired sensitivities. This can reduce the computation time for the objective function gradient dramatically.

However, there are some limitations. In NMPC we would like to specify some constraints on inputs and outputs. Gradient based techniques also need the gradient of these inequality constraints. Constraints on input variables can easily be added, but output constraints is more difficult to handle by using the adjoint method. That is, since outputs are coupled through nonlinear dynamics from the inputs, we also need the sensitivities from inputs to states which we avoided computing by using adjoints.

Some people have posed possible solutions to this problem. One approach suggest using inexact constraint jacobians with an update scheme and a modified gradient. The problem with this method is that we are unable to say something about the approximation error on line.

Another approach suggests lumping of the constrains into one constraint where the jacobian can be obtained by one adjoint simulation. The problem with this method is that by lumping the constraints, the optimization can produce an infeasible search direction. To account for this, one needs some type of recursive update scheme for modifying the control variables in order to produce a feasible direction. This method has been used for reservoir optimization where there for example can be some constraint on the total amount of liquid produced. In this case there is possible to find some way of reducing the producer valves in order to satisfy the constraints. In general applications there may be more difficult to find such a scheme to modify the controls in order to stay feasible.