1 Optimal control with ODE constraints

Lets try to derive the adjoint equation and the gradient, when we let $E(y, u)$ be the following ODE:

$$\begin{cases} y'(t) = \alpha y(t) + u(t), \ t \in (0, T) \\ y(0) = y_0 \end{cases}$$

We also choose the functional to be

$$J(y, u) = \frac{1}{2} \int_0^T u(t)^2 dt + \frac{1}{2}(y(T) - y^T)^2$$

Before we calculate the different terms in the gradient, we want to fit our ODE into an expression E. we do this by "moving" the initial condition into the equation:

$$E(y, u) = y' - \alpha y - u + \delta_0(y - y_0)$$

Here the δ_0 means evaluation at 0. Now lets find E_u, E_y, $\langle J_u(u), s\rangle$ and J_y with respect to our E and J.

$$\begin{aligned} E_u(y(u), u) &= -1 \\ E_y(y(u), u) &= \frac{\partial}{\partial t} - \alpha + \delta_0 \text{ ,where } \delta_0 \text{ is evaluation at } 0 \\ \langle J_u(u), s\rangle &= \int_0^T u(t)s(t)dt \end{aligned}$$

Lets be more thorough with J_y, which is the right hand side in the adjoint equation.

$$\begin{aligned} J_y(y(u), u) &= \frac{\partial}{\partial y}(\frac{1}{2}\int_0^T u^2 dt + \frac{1}{2}(y(T) - y^T)^2) \\ &= \frac{\partial}{\partial y}\frac{1}{2}(y(T) - y^T)^2 \\ &= \frac{\partial}{\partial y}\frac{1}{2}(\int_0^T \delta_T(y - y^T)dt)^2 \\ &= \delta_T \int_0^T \delta_T(y(t) - y^T)dt \\ &= \delta_T(y(T) - y^T) = L \end{aligned}$$

We have $E_y(y(u), u) = \frac{\partial}{\partial t} - \alpha + \delta_0$, but for the adjoint equation we need to find E_y^*. To derive the adjoint of E_y, we will apply it to a function v and then take the L^2 inner product with another function w. The next step is then to try to "move" the operator E_y from v to w. As becomes clear below, partial integration is the main trick to achieve this:

$$\begin{aligned}
\langle E_y v, w\rangle &= \int_0^T (v'(t) - \alpha v(t) + \delta_0 v(t))w(t)dt \\
&= \int_0^T v'(t)w(t)dt - \alpha \int_0^T v(t)w(t)dt + v(0)w(0) \\
&= -\int_0^T v(t)w'(t)dt + v(t)w(t)|_0^T - \alpha\langle v, w\rangle + v(0)w(0) \\
&= -\int_0^T v(t)w'(t)dt - \alpha\langle v, w\rangle + v(T)w(T) \\
&= \langle v, Pw\rangle
\end{aligned}$$

Where $P = -\frac{\partial}{\partial t} - \alpha + \delta_T$. This means that $E_y^* = P$, and we now have the left hand side in the adjoint equation. The right hand side is $J_y(y(u), u) = L$, which we have already found. If we write the adjoint equation on variational form it will look like this: $\langle Pp, w\rangle = \langle L, w\rangle$. To get back to standard ODE form, we can do some manipulation:

$$\begin{aligned}
\langle -p' - \alpha p + \delta_T p, w\rangle &= \langle \delta_T(y(T) - y^T), w\rangle \\
\langle -p' - \alpha p, w\rangle &= \langle \delta_T(y(T) - y^T - p), w\rangle
\end{aligned}$$

The right hand side is point evaluation at $t = T$, while the left hand side is an expression for all t. This finally gives us our adjoint equation:

$$\begin{cases} -p'(t) = \alpha p(t) \\ p(T) = y(T) - y^T \end{cases}$$

This is a simple and easily solvable ODE.

Expression for the gradient

We now have all the ingredients for finding an expression for the gradient of $\hat{J}$. If we remember that $\langle \hat{J}'(u), s\rangle = \langle y'(u)^* J_y(u), s\rangle + \langle J_u(u), s\rangle$, and all the different expressions for all the terms we calculated, we find:

$$\begin{aligned}
\langle \hat{J}'(u), s\rangle &= \langle y'(u)^* J_y(u), s\rangle + \langle J_u(u), s\rangle \\
&= \langle -E_u^* p, s\rangle + \langle J_u(u), s\rangle \\
&= \langle -(-1)^* p, s\rangle + \langle u, s\rangle \\
&= \langle p + u, s\rangle \\
&= \int_0^T (p(t) + u(t))s(t)dt
\end{aligned}$$

Note that the adjoint of a constant is just the constant itself.