Adjoint Equation

Looking at an optimal control problem

$$\min_{y,u} J(y,u) \text{ subject to } E(y,u) = 0$$

Where

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

and

$$E(y,u) = y' - \alpha y - u \tag{2}$$
$$y(0) = y_0 \tag{3}$$

Differentiating J is required for solving the problem. To do this we reduce (1) to $\hat{J}(u) = J(y(u), u)$ and computing its gradient in direction s. Will use the notation: $\langle \hat{J}'(u), s\rangle$ for the gradient.

$$\langle \hat{J}'(u), s\rangle = \langle \frac{\partial J(y(u),u)}{\partial u}, s\rangle \tag{4}$$
$$= \langle \frac{\partial y(u)}{\partial u}^* J_y(y(u),u), s\rangle + \langle J_u(y(u),u), s\rangle \tag{5}$$
$$= \langle y'(u)^* J_y(u), s\rangle + \langle J_u(u), s\rangle \tag{6}$$

Here $\langle \cdot, \cdot \rangle$ is the L^2 inner product. The difficult term in (6) is $y'(u)^*$,so lets first differentiate $E(y(u),u) = 0$ with respect to u, and try to find an expression for $y'(u)^*$:

$$\frac{\partial}{\partial u} E(y(u),u) = 0 \Rightarrow E_y(y(u),u)y'(u) = -E_u(y(u),u) \tag{7}$$
$$\Rightarrow y'(u) = -E_y(y(u),u)^{-1}E_u(y(u),u) \tag{8}$$
$$\Rightarrow y'(u)^* = -E_u(y(u),u)^* E_y(y(u),u)^{-*} \tag{9}$$

This means that

$$y'(u)^* J_y(u) = -E_u(y(u),u)^* E_y(y(u),u)^{-*} J_y(u) = -E_u(y(u),u)\lambda \tag{10}$$

were λ is the solution of the adjoint equation

$$E_y(y(u),u)^*\lambda = J_y(u) \tag{11}$$

This again means that

$$\langle \hat{J}'(u), s\rangle = \langle -E_u(y(u),u)\lambda, s\rangle + \langle J_u(u), s\rangle \tag{12}$$

Before we derive the adjoint equation, lets find E_u, E_y and $\langle J_u(u), s\rangle$ with respect to (1),(2) and (3).

$$E_u(y(u), u) = -1 \tag{13}$$

$$E_y(y(u), u) = \frac{\partial}{\partial t} - \alpha + \delta_0 \text{ ,where } \delta_0 \text{ is evaluation at } 0 \tag{14}$$

$$\langle J_u(u), s\rangle = \int_0^T u(t)s(t)dt \tag{15}$$

The expression for E_u gives us:

$$\langle y'(u)^* J_y(u), s\rangle = \langle -E(y(u), u)^*\lambda, s\rangle = \int_0^T \lambda(t)s(t)dt \tag{16}$$

To derive the adjoint equation, we write the operator $E_y(y(u), u)$ applied to a function v on variational form, and try to find the adjoint of E_y:

$$\langle E_y v, w\rangle = \int_0^T (v'(t) - \alpha v(t) + \delta_0 v(t))w(t)dt \tag{17}$$

$$= \int_0^T v'(t)w(t)dt - \alpha \int_0^T v(t)w(t)dt + v(0)w(0) \tag{18}$$

$$= -\int_0^T v(t)w'(t)dt + v(t)w(t)|_0^T - \alpha\langle v, w\rangle + v(0)w(0) \tag{19}$$

$$= -\int_0^T v(t)w'(t)dt - \alpha\langle v, w\rangle + v(T)w(T) \tag{20}$$

$$= \langle v, Pw\rangle \tag{21}$$

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 of the equation is $J_y(y(u), u)$. Lets look closer at this term:

$$J_y(y(u), u) = \frac{\partial}{\partial y}(\frac{1}{2}\int_0^T u^2 dt + \frac{1}{2}(y(T) - y^T)^2) \tag{22}$$

$$= \frac{\partial}{\partial y}\frac{1}{2}(y(T) - y^T)^2 \tag{23}$$

$$= \frac{\partial}{\partial y}\frac{1}{2}(\int_0^T \delta_T(y - y^T)dt)^2 \tag{24}$$

$$= \delta_T \int_0^T \delta_T(y(t) - y^T)dt \tag{25}$$

$$= \delta_T(y(T) - y^T) = L \tag{26}$$

Our adjoint equation on variational form then becomes $\langle P\lambda, w\rangle = \langle L, w\rangle$, which we can write:

$$\langle -p' - \alpha p + \delta_T p, w\rangle = \langle \delta_T(y(T) - y^T), w\rangle \tag{27}$$

$$\langle -p' - \alpha p, w\rangle = \langle \delta_T(y(T) - y^T - p), w\rangle \tag{28}$$

This then gives us the ODE:

$$\begin{cases} -\lambda'(t) - \alpha\lambda(t) = 0 \\ \lambda(T) = y(T) - y^T \end{cases} \tag{29}$$

If we solve this equation and plug it into (12), we see that the gradient of J is

$$\langle \hat{J}'(u), s \rangle = \int_0^T (\lambda(t) + u(t))s(t)dt \tag{30}$$

Discretization

Let us discretize our interval $[0, T]$ using $N + 1$ points where

$$x_n = n\Delta t, \ i = 0, ..., N \quad \text{and} \tag{31}$$

$$\Delta t = \frac{T}{N} \tag{32}$$

We also let $y^n = y(x^n)$ and $u^n = u(x^n)$. The integrals in our functional and its gradient we evaluate using the trapezoidal rule, and we discretize our ODE $E(y, u)$ and the adjoint equation using the Backward Euler scheme. For $E(y, u)$ we get :

$$\frac{y^n - y^{n-1}}{\Delta t} = \alpha y^n + u^n \tag{33}$$

$$(1 - \alpha\Delta t)y^n = y^{n-1} + \Delta t u^n \tag{34}$$

$$y^n = \frac{y^{n-1} + \Delta t u^n}{1 - \alpha\Delta t} \tag{35}$$

Here the initial condition $y^0 = y_0$ is known. For the adjoint equation the initial condition is $\lambda^N = y^N - y^T$, and the Backward Euler scheme gives us:

$$-\frac{\lambda^n - \lambda^{n-1}}{\Delta t} - \alpha\lambda^n = 0 \tag{36}$$

$$\lambda^{n-1} - \lambda^n = \Delta t\alpha\lambda^n \tag{37}$$

$$\lambda^{n-1} = (1 + \Delta t\alpha)\lambda^n \tag{38}$$