# Penalty gradient

We have the penalized functional

$$J(y, u, \lambda) = \int_0^T u^2 dt + \frac{1}{2}(y_n(T) - y_T)^2 + \frac{\mu}{2}\sum_{i=1}^n (y_{i-1}(T_i) - \lambda_i)^2$$

Our state equation is divided into $n+1$ equations on each interval:

$$\frac{\partial}{\partial t} y_i = y_i + u$$

$$y_i(T_i) = \lambda_i$$

here $i = 0, ..., n$, $\lambda_0 = y_0$ and $0 = T_0 < T_1 < \cdots < T_n < T_{n+1} = T$. Want the gradient of the reduced functional:

$$\langle \hat{J}'(u, \lambda), (s, l) \rangle = \langle \frac{\partial y(u, \lambda)}{\partial(u, \lambda)}^* J_y(y(u, \lambda), u, \lambda), (s, l) \rangle + \langle J_u + J_\lambda, (s, l) \rangle$$

$$= \langle -(E_u + E_\lambda)p, (s, l) \rangle + \langle J_u + J_\lambda, (s, l) \rangle$$

Where p is the solution of the adjoint equation $E_y^* p = J_y$, and $E$ is our ODEs:

$$E^i(y, u, \lambda) = \frac{\partial}{\partial t} y_i - y_i - u + \delta_{T_i}(y_i - \lambda_i)$$

Lets differentiate $E$:

$$E_y^i = \frac{\partial}{\partial t} - 1 + \delta_{T_i}$$

$$E_u^i = -1$$

$$E_{\lambda_i}^i = -\delta_{T_i}$$

Lets differentiate $J$:

$$\langle J_u, s \rangle = \int_0^T us \ dt$$

$$J_{\lambda_i} = -\mu(y_{i-1}(T_i) - \lambda_i)$$

$$J_y = \delta_{T_{n+1}}(y_n(T_{n+1}) - y_T) + \mu \sum_{i=1}^n \delta_{T_i}(y_{i-1}(T_i) - \lambda_i)$$

We also need $(E_y^i)^*$

$$\int_{T_i}^{T_{i+1}} E_y^i w \ v \ dt = \int_{T_i}^{T-i+1} (\frac{\partial}{\partial t} w - w) \ v \ dt + w(T_i)v(T_i)$$

$$= \int_{T_i}^{T_{i+1}} -(\frac{\partial}{\partial t} v + v) \ w \ dt + w(T_{i+1})v(T_{i+1})$$

$$= \int_{T_i}^{T_{i+1}} (-\frac{\partial}{\partial t} - 1 + \delta_{T_{i+1}}) v \ w \ dt$$

this means that $(E_y^i)^* = -\frac{\partial}{\partial t} - 1 + \delta_{T_{i+1}}$. This gives us the following expressions for the adjoint equations:

$i = n$ case:

$$-\frac{\partial}{\partial t} p_n = p_n$$

$$p_n(T_{n+1}) = y_n(T_{n+1}) - y_T$$

$i \neq n$ cases:

$$-\frac{\partial}{\partial t} p_i = p_i$$

$$p_i(T_{i+1}) = \mu(y_i(T_i + 1) - \lambda_{i+1})$$

Lets put everything into our expression for or gradient:

$$\begin{aligned}
\langle \hat{J}'(u, \lambda), (s, l) \rangle &= \langle -(E_u + E_\lambda)p, (s, l) \rangle + \langle J_u + J_\lambda, (s, l) \rangle \\
&= \langle (1 + \sum_{i=1}^{n} \delta_{T_i})p, (s, l) \rangle + \int_0^T us \ dt - \mu \sum_{i=1}^{n} (y_{i-1}(T_i) - \lambda_i)l_i \\
&= \int_0^T (u + p)s \ dt + \sum_{i=1}^{n} (p_i(T_i) - \mu(y_{i-1}(T_i) - \lambda_i))l_i \\
&= \int_0^T (u + p)s \ dt + \sum_{i=1}^{n} (p_i(T_i) - p_{i-1}(T_i))l_i
\end{aligned}$$