

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 solved separately on $n + 1$ intervals:

$$\begin{aligned} \frac{\partial}{\partial t} y_i &= y_i + u \text{ for } t \in [T_i, T_{i+1}] \\ y_i(T_i) &= \lambda_i \end{aligned}$$

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

$$\begin{aligned} \langle \hat{J}'(u, \lambda), (s, l) \rangle &= \left\langle \frac{\partial y(u, \lambda)}{\partial (u, \lambda)}^* J_y(y(u, \lambda), u, \lambda), (s, l) \right\rangle + \langle J_u + J_\lambda, (s, l) \rangle \\ &= \langle -(E_u + E_\lambda)p, (s, l) \rangle + \langle J_u + J_\lambda, (s, l) \rangle \end{aligned}$$

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 :

$$\begin{aligned} E_y^i &= \frac{\partial}{\partial t} - 1 + \delta_{T_i} \\ E_u^i &= -1 \\ E_{\lambda_i}^i &= -\delta_{T_i} \end{aligned}$$

Lets differentiate J :

$$\begin{aligned} \langle J_u, s \rangle &= \int_0^T u s \, dt \\ J_{\lambda_i} &= -\mu(y_{i-1}(T_i) - \lambda_i) \\ J_y &= \delta_{T_{n+1}}(y_n(T_{n+1}) - y_T) + \mu \sum_{i=1}^n \delta_{T_i}(y_{i-1}(T_i) - \lambda_i) \end{aligned}$$

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

$$\begin{aligned} \int_{T_i}^{T_{i+1}} E_y^i w \, v \, dt &= \int_{T_i}^{T_{i+1}} \left(\frac{\partial}{\partial t} w - w \right) v \, dt + w(T_i)v(T_i) \\ &= \int_{T_i}^{T_{i+1}} -\left(\frac{\partial}{\partial t} v + v \right) w \, dt + w(T_{i+1})v(T_{i+1}) \\ &= \int_{T_i}^{T_{i+1}} \left(-\frac{\partial}{\partial t} - 1 + \delta_{T_{i+1}} \right) v \, w \, dt \end{aligned}$$

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:

$$\begin{aligned} -\frac{\partial}{\partial t} p_n &= p_n \\ p_n(T_{n+1}) &= y_n(T_{n+1}) - y_T \end{aligned}$$

$i \neq n$ cases:

$$\begin{aligned} -\frac{\partial}{\partial t} p_i &= p_i \\ p_i(T_{i+1}) &= \mu(y_i(T_i + 1) - \lambda_{i+1}) \end{aligned}$$

Lets put everything into our expression for or gradient:

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