Adjoint sensitivity analysis for ODEs and DAEs

November 7, 2013

Abstract

The focus of this summer research work is on ODE/DAE constrained optimization of the electrical power grid transmission network. An adjoint sensitivity analysis is used to solve the dynamics constrained optimization problem. We present this adjoint sensitivity formulation, and preliminary results, for a test ODE constrained optimization problem of maximizing the mechanical input power of an electrical generator subject to limiting its mechanical angle oscillation. In addition, we present a generalization of this formulation for optimizing the transmission network cost subject to DAE constraints.

1 Problem Definition: ODEs

Consider a model represented by the Ordinary Differential Equation (ODE):

$$\mathbf{x}' = f(t, \mathbf{x}, \theta), \quad t_0 \leq t \leq t_F, \quad \mathbf{x}\,(t_0) = G(\theta)$$

where θ is the vector of parameters. The cost function we seek to minimize is:

$$\mathcal{J}\,(\theta) = a + b^T \theta + \frac{1}{2}\theta^T \cdot \mathcal{C}^T \theta\,.$$

Here a is a scalar. Also,

$$\theta \in \mathbb{R}^p, \quad \mathbf{x} \in \mathbb{R}^n, \quad b \in \mathbb{R}^p \quad and \quad \mathcal{C} \in \mathbb{R}^{p\times p}.$$

1.1 The Lagrangian

Given the Lagrangian:

$$\mathcal{L} = a + b^T \theta + \frac{1}{2}\theta^T \cdot \mathcal{C}^T \theta - \int_{t_0}^{t_F} \lambda^T(t) \cdot (\mathbf{x}' - f(t, \mathbf{x}, \theta))\ \mathrm{d}t$$

Taking variations we get:

$$
\begin{aligned}
\delta\mathcal{L} \quad = \quad & b \cdot \delta\theta + \left(\mathcal{C}^T \theta\right) \cdot \delta\theta \\
& + \int\limits_{t_0}^{t_F} \left(\left(f_\theta^T \left(\mathbf{x}(t), \theta\right) \cdot \lambda \right) \cdot \delta\theta + \left(f_{\mathbf{x}}^T \left(t, \mathbf{x}, \theta\right) \cdot \lambda \right) \cdot \delta\mathbf{x} \right) \mathrm{d}t \\
& - \int\limits_{t_0}^{t_F} \delta\lambda^T \cdot \left(\mathbf{x}' - f(t, \mathbf{x}, \theta)\right) \mathrm{d}t - \int\limits_{t_0}^{t_F} \lambda^T \cdot \left(\delta\mathbf{x}'\right) \mathrm{d}t
\end{aligned} \tag{1}
$$

Further, by performing integration by parts we get:

$$
\begin{aligned}
- \int\limits_{t_0}^{t_F} \lambda^T \cdot \left(\delta\mathbf{x}'\right) \mathrm{d}t &= - \lambda^T \left(t\right) \cdot \delta\mathbf{x} \left(t\right) \Big|_{t_0}^{t_F} + \int\limits_{t_0}^{t_F} \left(\lambda'\right)^T \cdot \delta\mathbf{x} \left(t\right) \, \mathrm{d}t \\
&= - \lambda^T \left(t_F\right) \cdot \delta\mathbf{x} \left(t_F\right) + \lambda^T \left(t_0\right) \cdot \delta\mathbf{x} \left(t_0\right) \\
&\quad + \int\limits_{t_0}^{t_F} \left(\lambda'\right)^T \cdot \delta\mathbf{x} \left(t\right) \, \mathrm{d}t \\
&= - \lambda^T \left(t_F\right) \cdot \delta\mathbf{x} \left(t_F\right) + \lambda^T \left(t_0\right) \cdot \left(\mathbf{x}_\theta(t_0) \cdot \delta\theta\right) \\
&\quad + \int\limits_{t_0}^{t_F} \left(\lambda'\right)^T \cdot \delta\mathbf{x} \left(t\right) \, \mathrm{d}t
\end{aligned} \tag{2}
$$

Setting $\langle \mathcal{L}_{\mathbf{x}}, \delta\mathbf{x} \rangle = 0$ we get :

$$
\begin{aligned}
& \left\langle \delta\mathbf{x} \left(t_F\right), -\lambda^T \left(t_F\right) \right\rangle + \left\langle \lambda', \delta\mathbf{x} \left(t\right) \right\rangle \\
& \qquad + \left\langle \left(f_{\mathbf{x}}^T \left(t, \mathbf{x}, \theta\right) \lambda \right), \delta\mathbf{x} \right\rangle \quad = \quad 0 \,.
\end{aligned} \tag{3}
$$

Equation (3) gives us the adjoint ODE:

$$
\lambda' = - f_{\mathbf{x}}^T \left(t, \mathbf{x}, \theta\right) \cdot \lambda, \qquad \lambda \left(t_F\right) = 0. \tag{4}
$$

The above adjoint ODE ((4)) has to be solved backwards. It requires the values of $\mathbf{x}$ at different times, which can be obtained by performing a forward run and checkpointing the solution. Furthermore, substituting (4) in (1) we obtain:

$$
\begin{aligned}
\langle \mathcal{L}_\theta, \delta\theta \rangle \quad = \quad & \left\langle f_\theta^T (t, \mathbf{x}, \theta) \cdot \lambda, \, \delta\theta \right\rangle + \left\langle b + \mathcal{C} \cdot \theta, \, \delta\theta \right\rangle \\
& + \left\langle G_\theta^T (t_0) \cdot \lambda(t_0), \, \delta\theta \right\rangle \,.
\end{aligned} \tag{5}
$$

Equation (5) gives us the expression for the gradient:

$$
\nabla_\theta \mathcal{J} \quad = \quad \int\limits_{t_0}^{t_F} \left(f_\theta^T (t, \mathbf{x}, \theta) \cdot \lambda \right) \mathrm{d}t + b + \mathcal{C} \cdot \theta + G_\theta^T (t_0) \cdot \lambda(t_0) \,. \tag{6}
$$

2 Example:ODEs

Consider a model represented by the following system of ODEs:

$$\begin{aligned}
\frac{\mathrm{d}\phi}{\mathrm{d}t} &= \omega_B\left(\omega-\omega_S\right)\,, \\
\frac{\mathrm{d}\omega}{\mathrm{d}t} &= \frac{\omega_S}{2H}\left(p_m-p_{max}\sin\left(\phi\right)-D\left(\omega-\omega_S\right)\right), \quad t_0\leq t\leq t_F\,.
\end{aligned}$$

Here ϕ is the phase angle and ω is the frequency. The objective is to maximize p_m, subject to the above ODEs as constraints and $\phi \leq \phi_S$ during all times. To begin with, we avoid the inequality constraint by adding an extra integral term, which, forces ϕ to be below ϕ_S. The cost function then becomes:

$$\Psi\left(p_m,\phi\right)=-p_m+c\int_{t_0}^{t_F}\left(\left(\phi-\phi_S\right)_+\right)^4\,\mathrm{d}t\,. \tag{7}$$

Here p_m is a parameter which can be tuned appropriately. The parameter controls the initial conditions for the ODE. The initial conditions is given by:

$$\begin{aligned}
\phi\left(t_0\right) &= \sin^{-1}\left(\frac{p_m}{p_{max}}\right)\,, \\
\omega\left(t_0\right) &= 1\,.
\end{aligned}$$

We construct the Lagrangian and determine the adjoint ODE so as to obtain the discrete gradients for the above cost function:

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t}=-J^T\lambda-\begin{bmatrix}\frac{\mathrm{d}\left(\left(\phi-\phi_S\right)_+\right)^4}{\mathrm{d}t}\\ 0\end{bmatrix}, \quad t_F\geq t\geq t_0\,. \tag{8}$$

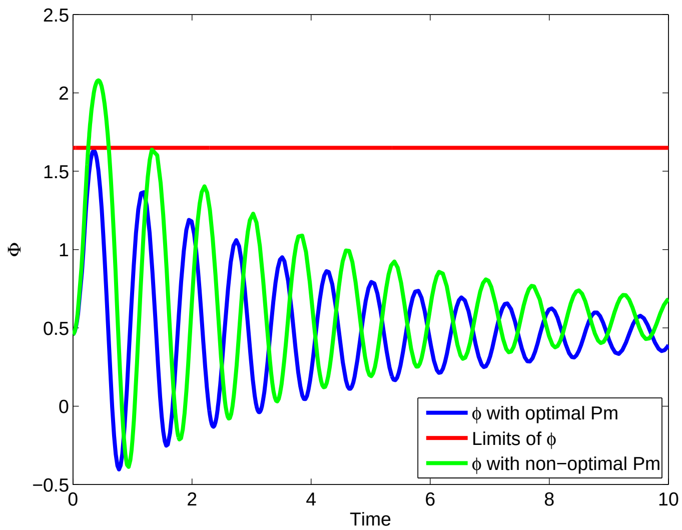
The initial conditions for (8) is given by: $\lambda\left(t_F\right)=\begin{bmatrix}0\\0\end{bmatrix}$.

Finally the gradient is given by:

$$\frac{\mathrm{d}\Psi}{\mathrm{d}p_m}=-1+\frac{\mathrm{d}\left[\phi\left(t_0\right)\omega\left(t_0\right)\right]}{\mathrm{d}p_m}\lambda\left(t_0\right)\,. \tag{9}$$

To obtain the integral in (7), we treat it as a quadrature term during the forward model run and obtain the cost function. To obtain the derivative in (8) we use finite differences. It should be noted that this does not require any extra computations, we just reuse the values obtained during the cost function computations. Finally once we have the gradient, we perform the optimization using the Poblano toolbox [**?**]. We solve the ODE using the optimal p_m and the plots are shown in the figure 1. The following enumerates the values of the certain constants, tolerances and initial guess on p_m.

1. $c = 10000$.
2. Initial guess for p_m : 0
3. For solving the ODEs, we use ODE15s and use the default tolerances (which is 1e-6 for both AbsTol and RelTol). We checkpoint the forward solution after every 1e-3 seconds.



(a) Plots of Φ with optimal p_m

Figure 1: The blue line shows the evolution of the phase angle using the optimal value of p_m with time and the red line shows the maximum acceptable value of phase angle.

3 General Formulation

$$\begin{aligned}
\dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{y}_g, \mathbf{y}_l, \theta_g), \quad \mathbf{x} \in \mathbb{R}^{m \times N_g} \\
g_g(\mathbf{y}_g, \mathbf{y}_l, \mathbf{x}, \mathbf{Y}) &= 0, \quad \mathbf{y}_g \in \mathbb{R}^{m \times N_g}, \quad \mathbf{y}_l \in \mathbb{R}^{m \times N_l}. \\
g_l(\mathbf{y}_g, \mathbf{y}_l, \theta_l, \mathbf{Y}) &= 0.
\end{aligned}$$

The initial conditions can be obtained by setting the differential equations to be $\mathbf{0}$. The parameter $\mathbf{Y}$ changes in case of a failure, this changes the dynamics and deviates from the steady state. $\mathbf{Y}$ is given by:

$$\mathbf{Y} = \begin{cases} \mathbf{Y}_0, & t = 0 \\ \mathbf{Y}_1, & t \in (0, T_c] \\ \mathbf{Y}_2, & t \in (T_c, \infty) \end{cases}$$

The differential equations correspond to the generator equations. There are N_g generators and m differential equations are associated with each of the generator. The algebraic equations consist of the stator algebraic equations and the network equations. It should be noted that $\mathbf{y}_g$ represents the generator voltages and currents, similarly $\mathbf{y}_l$ represent the load voltages and currents, θ_g and θ_l represent the control parameters associated with the generator and loads respectively. $\mathbf{Y}$ is the admittance matrix. The values of θ_g can be controlled by tuning the knob and they have to be tuned in such a way that there are violations in $\mathbf{x}$. The next step is to construct the Lagrangian and evaluate the associated gradient.

4 Gradient Computation

Let the cost function be $\Psi(\mathbf{x}, \theta_g)$. We construct the following Lagrangian:

$$\mathcal{L} = \Psi - \int_{t_0}^{t_F} \lambda^T \cdot (\dot{\mathbf{x}} - f) \, \mathrm{d}t - \int_{t_0}^{t_F} \mu^T \cdot g^g \, \mathrm{d}t - \int_{t_0}^{t_F} \nu^T \cdot g^l \, \mathrm{d}t$$

Taking variations we get:

$$\begin{aligned}
\delta\mathcal{L} = {} & \Psi_{\mathbf{x}} \delta\mathbf{x} + \Psi_{\theta_g} \delta\theta_g - \int_{t_0}^{t_F} \lambda^T \cdot \left(\delta\dot{\mathbf{x}} - f_{\mathbf{x}} \delta\mathbf{x} - f_{y_g} \delta y_g - f_{y_l} \delta y_l\right) \mathrm{d}t \\
& - \int_{t_0}^{t_F} \mu^T \cdot \left(g^g_{y_g} \delta y_g + g^g_{y_l} \delta y_l + g^g_{\mathbf{x}} \cdot \delta\mathbf{x}\right) \mathrm{d}t - \int_{t_0}^{t_F} \nu^T \cdot \left(g^l_{y_g} \delta y_g + g^l_{y_l} \delta y_l\right) \mathrm{d}t
\end{aligned}$$

From the above set of equations, we obtain the following adjoint DAE:

$$\begin{aligned}
\frac{d\lambda}{dt} + f_{\mathbf{x}}^{T}\lambda - g_{\mathbf{x}}^{g\,T}\mu + \Psi_{\mathbf{x}} &= 0, \quad \lambda\left(t_F\right) = 0, \quad t_F \geq t \geq t_0 \\
f_{y_g}^{T}\lambda - g_{y_g}^{g\;T}\mu - g_{y_g}^{l\;\;T}\nu &= 0 \\
f_{y_l}^{T}\lambda - g_{y_l}^{g\;T}\mu - g_{y_l}^{l\;\;T}\nu &= 0\,,
\end{aligned}$$

and we obtain the gradient as:

$$\frac{\partial \Psi}{\partial \theta_g} = \Psi_{\theta_g} + \left(\frac{d\mathbf{x}_0}{d\theta_g}\right)^{T} \cdot \lambda\left(t_0\right)$$