Supplementary Information
Solving Nonlinear and Complex Optimal Control Problems via Multi-task Artificial Neural Networks

Ali Emami Kerdabadi[1] Alaeddin Malek[1*]

[1]Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, 14115-134, Iran
Email: mala@modares.ac.ir

Preliminaries

Definition 1. *A convex polyhedron $P \in \mathbb{R}^n$ is defined by a finite set of linear inequalities*

$$P = \{x \in \mathbb{R}^n | Ax \leq b\}, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$.

Definition 2. *Let $S \subset \mathbb{R}^n$ be a compact set and $x \in \mathbb{R}^n$. The tangent cone of S at x denoted by $\mathcal{T}_S(x)$ and given as follow*

$$\mathcal{T}_S(x) = \{y \in \mathbb{R}^n | \liminf_{h\to 0} \frac{dist(x+hy, S)}{h} = 0\}$$

where $dist(x, S) = \inf_{s\in S} \|x - s\|$

Definition 3. *Suppose $\dot{x} = f(x,t)$ is a dynamical system. A set $S \subset \mathbb{R}^n$ is a positive invariant set if $x(0) \in S$ implies that $x(t) \in S$, for all $t \geq 0$.*

Theorem 1. *assume that $S \in \mathbb{R}^n$ be a closed and convex set. Consider the system $\dot{x} = f(x,t)$, where f is a continuous mapping, has a globally unique solution for every initial point $x(0) \in S$. Then S is an invariant set for this system if and only if*

$$f(x,t) \in \mathcal{T}_S(x), \quad \forall x \in \partial S$$

where $\mathcal{T}_S(x)$ is the tangent cone of S at x.

*Corresponding author, Email Adress: mala@modares.ac.ir

Lemma 1. *Let a polyhedron P that given in (1), and $\mathcal{I}_x \neq \emptyset$ for all $x \in P$. Then P is an invariant set for the continuous system $\dot{x} = f(x,t)$ if and only if for every $x \in \partial P$ we have*

$$A_i^T f(x,t) \leq 0, \quad \forall i \in \mathcal{I}_x$$

where $\mathcal{I}_x$ denote the set of indices of the constraints in (1) which are active at x.

Corollary 1. *$\mathbb{R}^{10}_{+,0}$ is an invariant set for the dynamical system given by $\dot{x} = f(x,t)$ if and only if for any point $x \in \partial S$, we have*

$$f_i(x,t) \geq 0, \quad \forall i \in \mathcal{I}_x$$

Proof. The proof is a direct consequence of Theorem 1 with $A = -I_n$ and $b = 0$. □

Supplementary Note 1: Optimal control analysis of an SIR epidemic model

Recall that the basic SIR model is given by the following system of ordinary differential equations:

$$\begin{aligned} \frac{dS}{dt} &= b - \beta S(t)I(t) - dS(t) - u_1(t)S(t), \\ \frac{dI}{dt} &= \beta S(t)I(t) - u_2(t)I(t) - dI(t) - \alpha I(t), \\ \frac{dR}{dt} &= u_1(t)S(t) + u_2(t)I(t) - dR(t). \end{aligned} \tag{2}$$

In this model, $S(t)$ denotes the number of susceptible individuals at time t, $I(t)$ denotes the number of infectious individuals, and $R(t)$ denotes the number of recovered (or removed) individuals. The parameter b represents the recruitment rate into the population, β denotes the disease transmission rate, and d is the natural death rate. Furthermore, u_1 denotes the proportion of susceptible individuals that is vaccinated per unit time, u_2 denotes the proportion of infectious individuals that is treated per unit time, and α represents the disease-induced death rate.

Theorem 2. *Let $\mathbb{R}^3_{+,0} = \{(x_1,x_2,x_3) \in \mathbb{R}^3 | x_i \geq 0, i = 1,2,3\}$ represent non-negative quadrant of $\mathbb{R}^3$. then $\mathbb{R}^3_{+,0}$ is an invariant set with respect to the dynamical system (2).*

Proof. Let $E_i = \{(x_1,x_2,x_3) \in \mathbb{R}^3 | x_i = 0, x_j \geq 0, j = 1,2,3, \ j \neq i\}, i = 1,2,3,$ then the boundary of $\mathbb{R}^3_{+,0}$ can be written as

$$\partial \mathbb{R}^3_{+,0} = \bigcup_{i=1}^{3} E_i$$

Suppose $x = (S, I, R) \in \partial\mathbb{R}^3_{+,0}$, then for dynamical system (2) we have

$$
\begin{aligned}
&\text{if } x \in E_1, && \frac{dS}{dt} = b \geq 0 \\
&\text{if } x \in E_2, && \frac{dI}{dt} = 0 \geq 0 \\
&\text{if } x \in E_3, && \frac{dR}{dt} = u_1(t)S(t) + u_2(t)I(t) \geq 0,
\end{aligned}
$$

By using Corollary 1, the result follows immediately. □

Theorem 3. *The solutions of dynamical system (2) with the initial conditions satisfying the following inequalities*

$$S(0) \geq 0,\ I(0) \geq 0,\ R(0) \geq 0,$$

are bounded.

Proof. The total population $N(t) = S(t) + I(t) + R(t)$, it implies that

$$\frac{dN}{dt} = b - N(t)d - \alpha I(t) \leq b - N(t)d$$

Then $\limsup_{t\to\infty} \leq \frac{b}{d}$. Therefore all solutions are bounded □

The SIR optimal control problem is governed by

$$
\begin{aligned}
&\min_u \quad I(T) + \frac{1}{2}\int_0^T C_1 u_1(t)^2 + C_2 u_2(t)^2 \ dt \\
&\text{Subject to} \quad \begin{cases} \dot{S}(t) = & b - \beta S(t)I(t) - dS(t) - u_1(t)S(t), \\ \dot{I}(t) = & \beta S(t)I(t) - u_2(t)I(t) - dI(t) - \alpha I(t), \\ \dot{R}(t) = & u_1(t)S(t) + u_2(t)I(t) - dR(t). \end{cases}
\end{aligned} \tag{3}
$$

where $C1$ and C_2 are the weights for the control effort. T is the final time of the intervention period. The idea is to define the Hamiltonian H associated with the problem:

$$
\begin{aligned}
H = {} & C_1 u_1(t)^2 + C_2 u_2(t)^2 + \lambda_S(t)\Big[b - \beta S(t)I(t) - dS(t) - u_1(t)S(t)\Big] \\
& + \lambda_I(t)\Big[\beta S(t)I(t) - u_2(t)I(t) - dI(t) - \alpha I(t)\Big] + \lambda_R(t)\Big[u_1(t)S(t) + u_2(t)I(t) - dR(t)\Big],
\end{aligned}
$$

where $\lambda_S(t)$, $\lambda_I(t)$ and $\lambda_R(t)$ are the adjoint variables. Addition the adjoint variables satisfy

$$
\begin{aligned}
\dot{\lambda}_S(t) &= \lambda_S(t)\beta I(t) + \lambda_S(t)d + \lambda_S(t)u_1 - \lambda_I(t)\beta I(t) \\
\dot{\lambda}_I(t) &= \lambda_S(t)\beta S(t) - \lambda_I(t)\beta S(t) + \lambda_I(t)u_2 + \lambda_I(t)d + \lambda_I(t)\alpha \\
\dot{\lambda}_R(t) &= d\lambda_R(t)
\end{aligned}
$$

Supplementary Note 2: Optimal Control of the Kuramoto Model

The characterization of the optimal control follows from Pontryagin's Maximum Principle

$$H(t,x,\lambda,u) = \frac{1}{3^2}\Big((\sum_{j=1}^{3} sin(\theta_j))^2 + (\sum_{j=1}^{3} cos(\theta_j))^2\Big)$$
$$- \sum_{i\neq j} B_{ij}u_{ij} + \sum_{i=1}^{3} \lambda_i \Big(\omega_i + \frac{1}{3}\sum_{j\neq i} u_{ij} sin(\theta_j - \theta_i)\Big)$$

The adjoint variables for $i = 1, 2$ & 3 satisfy

$$\dot{\lambda}_i(t) = \frac{1}{3}\sum_{j\neq i}(\lambda_i u_{ij} - \lambda_j u_{ji}) cos(\theta_j - \theta_i) - \frac{2}{3^2}\sum_{j\neq i} sin(\theta_j - \theta_i)$$

where $\lambda_i(T) = 0$ for $i = 1, 2$ & 3.

Supplementary Note 3: proof

Lemma 2. *Let $f : S \subset \mathbb{R}^n \to \mathbb{R}_+$ is a convex and differentiable function and S is closed. Consider the unconstrained scalar-valued minimization problem*

$$v^* = argmin_{v\in\mathbb{R}^n} max(k_1\|v\|, k_2 F(x+v))$$

then, there are $k_1, k_2 > 0$ such that

$$\forall m > 0, \qquad \|f(x+v^*) - f(x)\| \leq m$$

Proof. Consider the unconstrained scalar-valued minimization problem

$$v(x) = argmin_{v\in\mathbb{R}^n} max(k_1\|v\|, k_2 F(x+v))$$

there is $v \neq 0$ such that $f(x+v) < f(x)$. In addition for $\exists\ k1, k2 > 0$ such that $v^* = argmin_{v\in\mathbb{R}^n} max(k_1\|v\|, k_2 F(x+v))$ and

$$k_1\|v^*\| = k_2 f(x+v^*) = max(k_1\|v^*\|, k_2 F(x+v^*))$$

f is Lipschitz continuous, then

$$\|f(x+v^*) - f(x)\| < L\|v^*\| < L\frac{k_2}{k1} f(x+v^*)$$

S is closed then

$$\exists\ M \qquad s.t. \qquad f(x+v^*) \leq M$$

then we have

$$\|f(x+v^*) - f(x)\| < L\|v^*\| \leq L\frac{k_2}{k1} f(x+v^*) \leq \frac{k_2}{k1} M$$

its sufficient that $k1 > \frac{M}{m} k_2$ □

Corollary 2. *Let* $f = \|\frac{\partial H}{\partial u}\|$. *Note that Lemma (2),* $\frac{\partial H}{\partial u}$ *is valid for neighborhood* $B_m(u)$ *and by considering lipschitz continuity of neural networks, the iterative updates of the neural networks based on algorithm 1 converge to the optimal solution.*

References

1. Nagumo, M. (1942). Über die lage der integralkurven gewöhnlicher differentialgleichungen. Proceedings of the Physico-Mathematical Society of Japan. 3rd Series, 24, 551-559.

2. Horváth, Z., Song, Y., & Terlaky, T. (2017). A novel unified approach to invariance conditions for a linear dynamical system. Applied Mathematics and Computation, 298, 351-367.