Reinforcement Learning Seminar

Boyeon Kim

Department of Mathematics, School of Mathematics and Computing
Mathematics
Yonsei University

January 13, 2023

SLIAR Optimal control

- Goal1 : Reproduce the SLIAR optimal control by uning PMP
- Method : PMP (Adjoint)

SLIAR Model

- SLIAR Model Equations and structure.

$$\begin{cases} S' &= -\beta(1-\sigma)S\Lambda - \nu S \\ L' &= \beta(1-\sigma)S\Lambda - \kappa L \\ I' &= p\kappa L - \alpha I - \tau I \\ A' &= (1-p)\kappa L - \eta A \end{cases} \qquad \text{with } \Lambda = \epsilon L + (1-q)I + \delta A$$

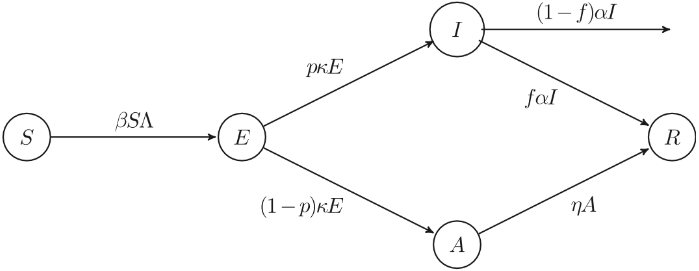
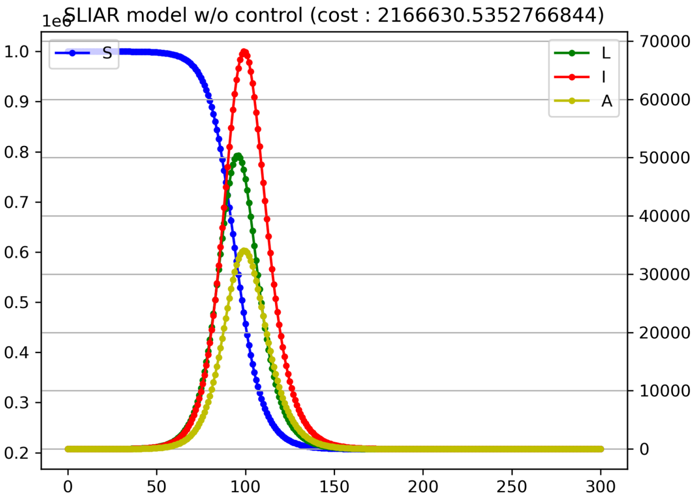


Fig. 1. SEIAR epidemic model.

SLIAR model parameters

Start	0	β	7.26e-07
End	300	σ	0
S0	1e06	κ	0.526
L0	0	α	0.224
I0	1	τ_{max}	0.05
A0	0	ν_{max}	0.01
R_0	1.9847	ϵ	0.224
P	1	q	0.5
Q	1e06	p	0.667
R	1e06	δ	1
W	0		

SLIAR w/o constrol



SLIAR optimal control

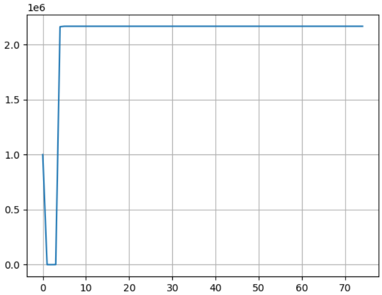
$$\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$$

subject to

$$\begin{cases} S' &= -\beta(1-\sigma)S\Lambda - \nu S \\ L' &= \beta(1-\sigma)S\Lambda - \kappa L \\ I' &= p\kappa L - \alpha I - \tau I \\ A' &= (1-p)\kappa L - \eta A \end{cases} \qquad with \quad \Lambda = \epsilon L + (1-q)I + \delta A$$

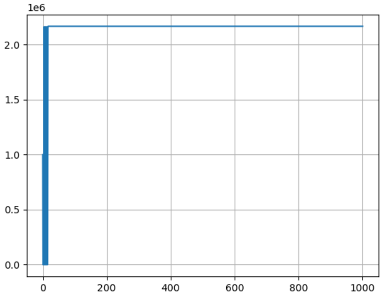
SLIAR optimal control

- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : PMP(Adjoint)
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, learning rate : 0.1



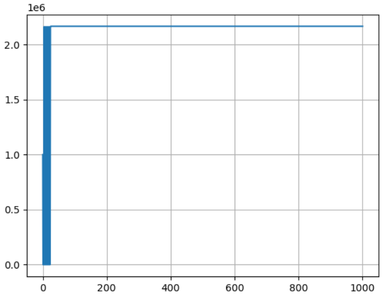
SLIAR optimal control

- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : PMP(Adjoint)
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, learning rate : 0.001



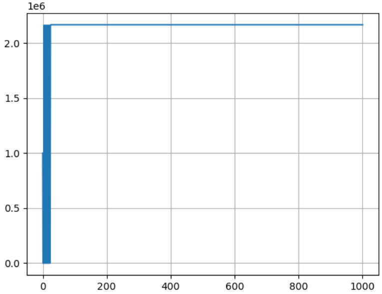
SLIAR optimal control

- $\min_{u\in\mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : PMP(Adjoint)
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, learning rate : 1e-04



SLIAR optimal control

- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : PMP(Adjoint)
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, learning rate : 1e-06

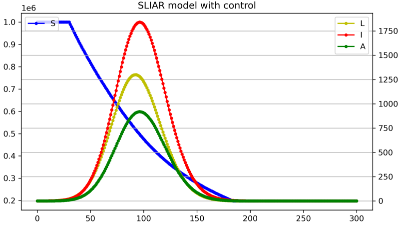


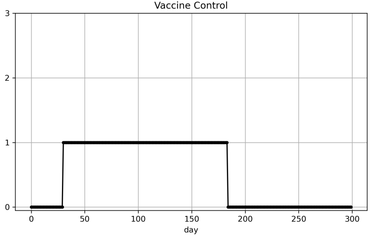
SLIAR Optimal control

- Goal2 : 2-constraint optimal control
- Method : DQN

SLIAR optimal control

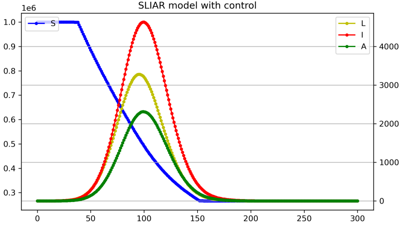
- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : DQN
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, $\tau_{max} = 0.05$,iteration : 2000

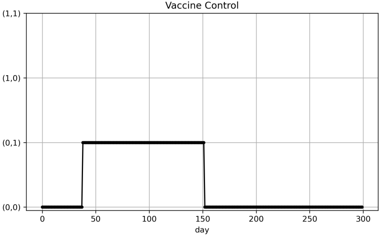




SLIAR optimal control

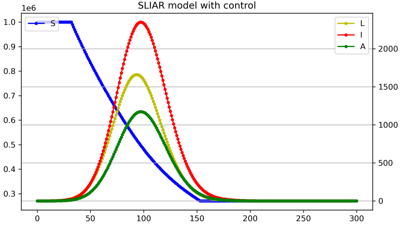
- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : DQN
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, $\tau_{max} = 0.05$,iteration : 5000

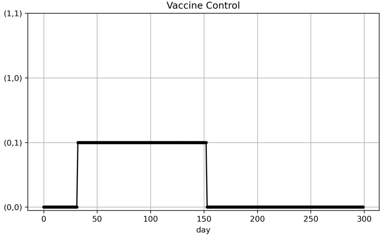




SLIAR optimal control

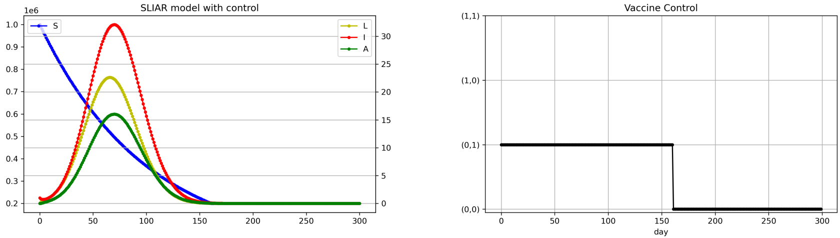
- $\min_{u\in\mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : DQN
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, $\tau_{max} = 0.05$,iteration : 7000





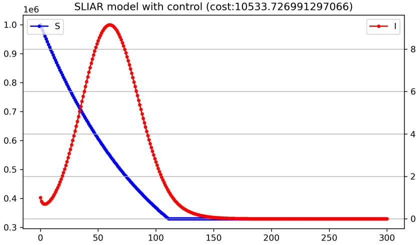
SLIAR optimal control

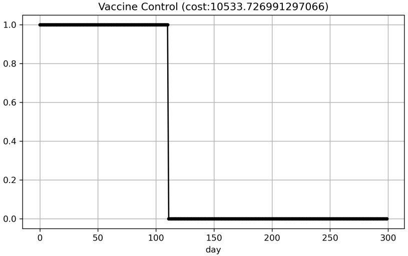
- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : DQN
- P = 1, Q = 1E6, R = 1E6, $\nu_{max} = 0.01$, $\tau_{max} = 0.05$,iteration : 10000



1-constraint result of DQN

- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : DQN
- P = 1, Q = 1e6, R = 1e6 ν_{max} = 0.01, iteration : 10,000





1-constraint result of DQN

- $\min_{u \in \mathcal{U}_{ad}} \int_0^T PI(t) + Q\nu^2(t) + R\tau^2(t) + W\sigma^2(t)dt$
- Method : DQN
- P = 1, Q = 1e6, R = 1e6τ_{max} = 0.05, iteration : 10,000