Learning optimal actions from experience

a neural approach

Corrado Possieri

Istituto di Analisi dei Sistemi ed Informatica “A. Ruberti”,
Consiglio Nazionale delle Ricerche, 00185 Roma, Italy

June 16, 2020

Finite-horizon optimization problem

Consider the discrete-time system

$$x_{k+1} = f_k(x_k, u_k),$$

and the cost function

$$J(x_0, u_0, \ldots, u_{N-1}) = \sum_{k=0}^{N} \left(\|x_k - x_k^\star\|^2 + \|u_k - u_k^\star\|^2 \right),$$

where $x_k^\star$ and $u_k^\star$ are reference trajectories and inputs.

Let U_k be the (convex) set of admissible inputs at time k.

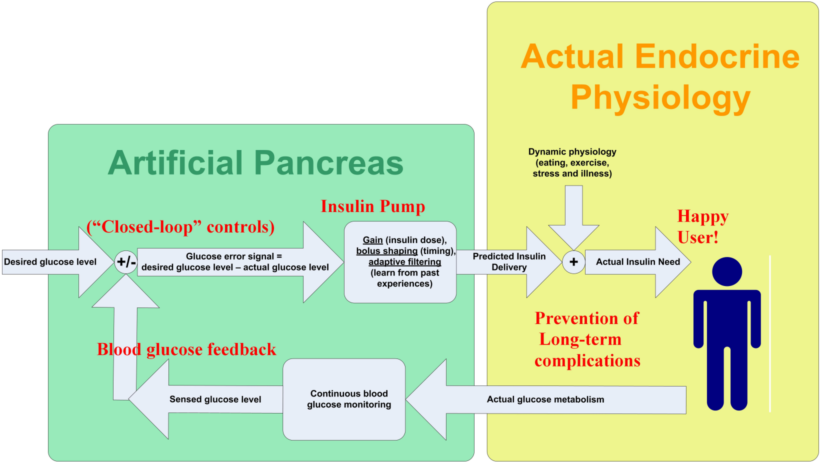
Optimal solution

Find a sequence $\{u_0^\star, \ldots, u_{N-1}^\star\}$ that minimizes J,

$$J^\star(x_0) := \min_{u_k \in U_k,\, k=0,\ldots,N-1} J(x_0, u_0, \ldots, u_{N-1}).$$

An example

Artificial pancreas



Classical solution

The classical solution to the problem consists in solving

$$\left|\begin{array}{rl} \min & \sum_{k=0}^{N} \left(\|x_k - x_k^\star\|^2 + \|u_k - u_k^\star\|^2\right), \\ \text{with} & x_{k+1} = f_k(x_k, u_k), \quad k = 0, \ldots, N-1, \\ & u_k \in U_k, \quad k = 0, \ldots, N-1, \end{array}\right.$$

feed the system with u_0 and repeat in a rolling horizon fashion.

Main drawbacks:

Curse of modeling: availability of an accurate model;

Curse of dimensionality: computational cost.

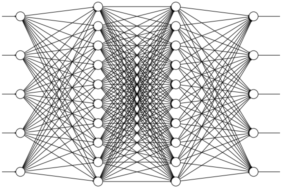
Possible solution to the modeling issue

Use a feedforward **neural network**.

Universal approximation theorem

For each real-valued continuous function $g(\cdot)$ defined on a compact, nonempty set $\mathcal{K} \subset \mathbb{R}^n$, and for each $\varepsilon > 0$ there exists a feedforward neural network $\phi(\cdot)$ such that

$$|\phi(x) - g(x)| \leq \varepsilon, \text{forall } x \in \mathcal{K}.$$



Use data to approximate

$$x_0, u_0, \ldots, u_{N-1} \xmapsto{\phi} J(x_0, u_0, \ldots, u_{N-1}).$$

The computational issue persists

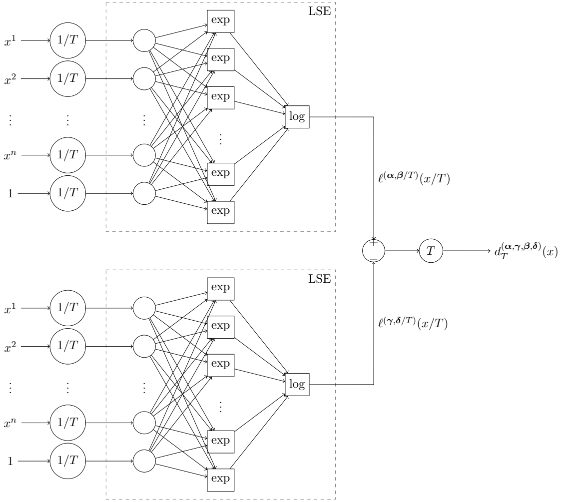
The critical issue of solving

$$\left|\begin{array}{ll} \min & \phi(x_0, u_0, \ldots, u_{N-1}), \\ \text{with} & u_k \in U_k, \quad k = 0, \ldots, N-1, \end{array}\right.$$

is that it is a **non-convex** optimization problem:

- extremely large computation times;
- sub-optimality of the solution;
- curse of dimensionality;
- local solutions.

DLSE_T neural networks



Universal approximation result

Functions in DLSE_T are[1]

- smooth;
- in a difference of **convex** form.

Universal approximation

Given a continuous $g(\cdot)$ defined on a compact convex set $\mathcal{K} \subset \mathbb{R}^n$, for all $\varepsilon > 0$ there exist $T > 0$ and $d_T \in \mathrm{DLSE}_T$ such that

$$|d_T(x) - g(x)| \leqslant \varepsilon, \quad \text{forall } x \in \mathcal{K}.$$

We can still use data to approximate

$$x_0, u_0, \ldots, u_{N-1} \stackrel{d_T}{\longmapsto} J(x_0, u_0, \ldots, u_{N-1}).$$

[1] G. C. Calafiore, S. Gaubert, and C. Possieri, "A Universal Approximation Result for Difference of log-sum-exp Neural Networks.", IEEE Transactions on Neural Networks and Learning Systems, 2020.

DLSE_T networks are easily trainable

A DLSE_T network is a function of the parameters $\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\delta}$.

We can compute the gradients

$$\nabla_{\alpha^{(i)}} d_T^{(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta},\boldsymbol{\delta})}(x) = \frac{\exp(\langle \alpha^{(i)}, x/T\rangle + \beta_i/T)\, x}{\sum_{k=1}^K \exp(\langle \alpha^{(k)}, x/T\rangle + \beta_k/T)},$$

$$\nabla_{\beta_i} d_T^{(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta},\boldsymbol{\delta})}(x) = \frac{\exp(\langle \alpha^{(i)}, x/T\rangle + \beta_i/T)}{\sum_{k=1}^K \exp(\langle \alpha^{(k)}, x/T\rangle + \beta_k/T)},$$

$$\nabla_{\boldsymbol{\gamma}^{(i)}} d_T^{(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta},\boldsymbol{\delta})}(x) = -\frac{\exp(\langle \boldsymbol{\gamma}^{(i)}, x/T\rangle + \delta_i/T)\, x}{\sum_{k=1}^K \exp(\langle \boldsymbol{\gamma}^{(k)}, x/T\rangle + \delta_k/T)},$$

$$\nabla_{\delta_i} d_T^{(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\beta},\boldsymbol{\delta})}(x) = -\frac{\exp(\langle \boldsymbol{\gamma}^{(i)}, x/T\rangle + \delta_i/T)}{\sum_{k=1}^K \exp(\langle \boldsymbol{\gamma}^{(k)}, x/T\rangle + \delta_k/T)},$$

and use classical training algorithms to train these networks.

DLSE_T network are easily optimizable

Input: functions $g_T = f_T^{(\boldsymbol{\alpha},\boldsymbol{\beta})}$ and $h_T = f_T^{(\boldsymbol{\gamma},\boldsymbol{\delta})}$ in LSE_T and a convex compact set $\mathcal{K}$.

Output: a candidate optimal solution $\hat{x}^\star$ to the problem

$$\min_{x\in\mathcal{K}} (g_T(x) - h_T(x)).$$

1: pick initial point $\chi^{(0)} \in \mathcal{K}$,
2: **for** $\varkappa \in \mathbb{N}$ **do**
3: let $v^{(\varkappa)} = \frac{\sum_{k=1}^K \exp(\langle \boldsymbol{\gamma}^{(k)}, \chi^{(\varkappa)}/T\rangle + \delta_k/T)\,\boldsymbol{\gamma}^{(k)}}{\sum_{k=1}^K \exp(\langle \boldsymbol{\gamma}^{(k)}, \chi^{(\varkappa)}/T\rangle + \delta_k/T)}$,
4: let $\chi^{(\varkappa+1)} = \arg\min_{x\in\mathcal{K}}\{g_T(x) - \langle x, v^{(\varkappa)}\rangle\}$,
5: **if** $\frac{\|\chi^{(\varkappa+1)} - \chi^{(\varkappa)}\|}{1+\|\chi^{(\varkappa)}\|}$ is smaller than a tolerance **then**
6: **return** $\hat{x}^\star = \chi^{(\varkappa+1)}$.

Sub-optimality of the solution

Letting $\hat{u}$ be the solution to

$$\left|\begin{array}{ll} \min & d_T(x_0, u_0, \ldots, u_{N-1}), \\ \text{with} & u_k \in U_k, \quad k = 0, \ldots, N-1, \end{array}\right.$$

and letting u be the solution to

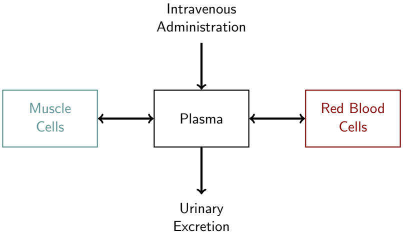
$$\left|\begin{array}{ll} \min & \sum_{k=0}^{N} \left(\|x_k - x_k^\star\|^2 + \|u_k - u_k^\star\|^2 \right), \\ \text{with} & x_{k+1} = f_k(x_k, u_k), \quad k = 0, \ldots, N-1, \\ & u_k \in U_k, \quad k = 0, \ldots, N-1, \end{array}\right.$$

we have that

$$|\hat{u} - u| < \epsilon, \qquad \text{for all } x_0 \in \mathcal{X}.$$

Lithium ion distribution

Consider the model for the distribution of Lithium ions in humans.



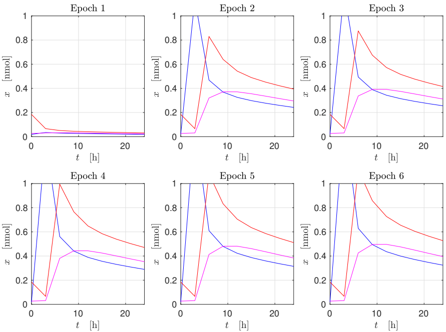
The objective is to let the concentration levels C satisfy

$$\begin{aligned} 0.4\,\mathrm{nmol/L} &\leqslant \mathrm{C}_{\mathrm{plasma}} \leqslant 0.6\,\mathrm{nmol/L}, \\ 0.6\,\mathrm{nmol/L} &\leqslant \mathrm{C}_{\mathrm{RBC}} \leqslant 0.9\,\mathrm{nmol/L}, \\ 0.5\,\mathrm{nmol/L} &\leqslant \mathrm{C}_{\mathrm{muscle}} \leqslant 0.8\,\mathrm{nmol/L}. \end{aligned}$$

Test of the LSE_T approach

Table: Average costs

Epoch	LSE networks
1	3.9925
2	2.4126
3	2.2527
4	1.9026
5	1.6736
6	1.6430



Proposal

Do you have problems with data, no model, and an intrinsic optimization to be performed?

Mail to: corrado.possieri@iasi.cnr.it.